

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 19, 2002, 16:02:26 ; Search time 28.5982 Seconds
(without alignments)
55.932 Million cell updates/sec

Title: US-08-653-294C-36

Perfect score: 58

Sequence: 1 YRLAIRIALRY 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 473727

Minimum DB seq length: 0

Maximum DB seq length: 60

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_101002:*

- 1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
- 23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	58	100.0	12	16 AAR95429	HLA-B2702 84-79-84
2	58	100.0	12	19 AAW33798	Peptide B2702.84-7
3	58	100.0	12	19 AAW33799	Immunomodulating d
4	44	75.9	20	16 AAR95428	HLA-B2702 84-75-84
5	44	75.9	20	16 AAR92907	HLA-B2702 CTL modu
6	44	75.9	20	19 AAW33778	Immunomodulating d
7	39	67.2	20	16 AAR92908	HLA-B2702 CTL modu
8	39	67.2	20	16 AAR92908	HLA-B2702 CTL modu
9	39	67.2	20	19 AAW33791	Peptide B2702.84-7
10	39	67.2	20	19 AAW33792	Peptide B2702.84-7

11	37	63.8	20	16 AAR95430	HLA-B2702 84-75T/7
12	35	60.3	18	16 AAR71429	Human MHC 1 alpha
13	35	60.3	46	22 AAM83083	Human immune/haema
14	34.5	59.5	20	16 AAR92911	HLA-B2702 CTL modu
15	34.5	59.5	20	19 AAW33779	Immunomodulating d
16	34	58.6	20	16 AAR92910	HLA-B2702 CTL modu
17	34	58.6	20	19 AAW33793	Peptide B2702.84-7
18	32	55.2	40	18 AAW12037	Curvularia verrucu
19	30	51.7	15	23 AAU98851	E.Coli DNAJ 22 imm
20	30	51.7	15	23 AAU98862	Human HDJ1 immunog
21	30	51.7	33	22 AAM30767	Peptide #4804 enco
22	30	51.7	51	22 AAU22828	Human prostate can
23	30	51.7	51	22 AAM96134	Human reproductive
24	30	51.7	55	22 AAU39568	Propionibacterium
25	30	51.7	55	22 AAU42856	Propionibacterium
26	30	51.7	56	23 ABP09392	Human OREF protein
27	30	51.7	57	21 AAG01773	Human secreted pro
28	30	51.7	58	22 AAU56052	Propionibacterium
29	29.5	50.9	12	22 AAB49012	Propionibacterium
30	29.5	50.9	21	22 AAB88757	Human interleukin-
31	29	50.0	6	19 AAW33780	Peptide #1 used in
32	29	50.0	6	19 AAW33781	Peptide #2 used in
33	29	50.0	6	19 AAW47263	Immunomodulatory p
34	29	50.0	6	19 AAW47261	Immunomodulatory p
35	29	50.0	9	22 AAY72479	Immunosuppressive
36	29	50.0	9	22 AAB59450	Human Class I HLA-
37	29	50.0	10	14 AAR41208	Peptide fragment o
38	29	50.0	10	16 AAR95413	Alphal-helix of HL
39	29	50.0	10	16 AAR95425	HLA-B2702.75-84(D)
40	29	50.0	10	16 AAR83062	HLA-B2702 CTL modu
41	29	50.0	10	16 AAR83094	HLA-B2702 CTL modu
42	29	50.0	10	17 AAW07512	T-cell modulating
43	29	50.0	10	17 AAW07513	T-cell modulating
44	29	50.0	10	19 AAW33784	Peptide B2702.75-8
45	29	50.0	10	19 AAW33787	Peptide B2702.75-8

ALIGNMENTS

RESULT 1

AAR95429
ID AAR95429 standard; peptide; 12 AA.

AC AAR95429;

DT 12-NOV-1996 (first entry)

DE HLA-B2702 84-79-84 palindrome.

XX HLA; p74; alphal-helix; human-leucocyte-associated antigen; inhibitor;
KW T-cell lyase; membrane protein; mammal; heat shock protein; Hsc70; APC;
KW B cell; calcium influx; cytotoxic T lymphocyte; CTL; differentiation;
KW cytolysis; antigen presenting cell.

OS Synthetic.

PN WO9513288-A1.

PD 18-MAY-1995.

PF 10-NOV-1994; 94WO-US12985.

XX 10-NOV-1993; 93US-0150493.

PA (STRD) UNIV LELAND STANFORD JUNIOR.

PI Clayberger C, Krensky AM;

DR WPI; 1995-194027/25.

PT Compens. comprising lymphoid surface membrane proteins - which may
inhibit cytolytic activity and differentiation of CTLs.

XX Example; Page 12; 29pp; English.
 PS AAR95413, and AAR95415-R95431 represent palindromes and fragments of
 XX human-leucocyte-associated antigens. This sequence represents the
 CC HLA-B*2702 84-79-84 palindromic. These sequences can be used to isolate
 CC the protein p74 from a T-cell lysate. p74 is a T-cell surface membrane
 CC protein associated with T-cell activation in mammalian T-cells, and is
 CC also immunologically cross reactive with the heat shock protein Hsc70.
 CC p74 is found in a limited number of cell types, but is particularly
 CC expressed on B and T cells. p74 can be isolated by lysis of a suitable
 CC cell with an amphoteric detergent, and then passed through an affinity
 CC column containing a covalently bound HLA-B*2702 palindromic peptide.
 CC Compositions comprising the extracellular fragment of p74 combined with
 CC HLA-B*2702.60-84 (see AAR95416), induces calcium influx, and inhibits
 CC cytotoxic T lymphocyte (CTL) differentiation or cytotoxicity. Candidate
 CC compounds can be screened for their effect on the cytolytic activity of
 CC T-cells, by combining them with the extracellular portion of p74 and
 CC determining the amount of binding between the candidate compound and
 CC p74. Modulation of CTL activity can be inhibited in a cellular
 CC composition containing T-cells and antigen presenting cells (APCs), by
 CC adding to the mix the extracellular portion of p74, in an amount
 CC sufficient to compete with p74 for the binding of the p74 ligand.
 CC
 SQ Sequence 12 AA;
 Query Match 100.0%; Score 58; DB 16; Length 12;
 Best Local Similarity 100.0%; Pred. No. 0.00026;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 YRLAIRRLRLRY 12
 Db 1 YRLAIRRLRLRY 12
 RESULT 2
 AAM33798
 ID AAM33798 standard; peptide; 12 AA.
 AC AAM33798;
 XX
 DT 19-UN-1998 (first entry)
 DE Peptide B2702.84-79/79-84 tested for immunomodulating activity.
 XX
 KW Immunomodulating dimer; immunosuppressant drug; CTL activation;
 KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;
 KW rejection.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 PN MO9744351-A1.
 XX
 PD 27-NOV-1997.
 XX
 PF 22-MAY-1997; 97WO-US08689.
 XX
 PR 24-MAY-1996; 96US-0653294.
 XX
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 XX
 PI Beulow R, Clayberger C, Krensky AM;
 XX
 DR WPI; 1998-086530/08.
 XX
 PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B
 PT alpha-1 domain, used for preventing rejection of transplants or
 PT treating autoimmune diseases
 XX
 PS Example 1; Page 19; 41pp; English.
 XX
 CC Peptides AAM33784-98 and AAM33778-9 were assayed for their

CC immunomodulating activity. A peptide-type compound or variant is claimed
 CC which has immunomodulating activity, including the N-terminal acylated
 CC and/or C-terminal amidated or esterified forms of up to 60 amino acids,
 CC where the peptide-type compound comprises the formula: A-B, where A, B =
 CC (R aa76-77L) (aa79-84) or (aa84-79) (aa77-76R); aa76 = E or V; aa77 =
 CC D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a hydrophobic or
 CC small amino acid; aa82 = R or L; aa83 = G or R; and aa represents amino
 CC acid. The sequence in the brackets may optionally be absent or truncated
 CC at any peptide type bond within the brackets. The compounds comprise
 CC amino acid sequences related to a Class I HLA-B alpha1 domain (positions
 CC 79-84). They can be used to inhibit cytotoxic T-lymphocytes (CTL) from
 CC undesirably attacking cells in a host or in vitro. They can also be
 CC used in combination with antigenic peptides or proteins of interest to
 CC activate CTLs. They can also inhibit the proliferation of T cells in
 CC response to anti-CD3. The peptide can be used for preventing rejection
 CC of transplants or for treating autoimmune diseases, e.g. diabetes,
 CC rheumatoid arthritis and lupus erythematosus. The products can also be
 CC used for detection and diagnosis.
 CC
 SQ Sequence 12 AA;
 Query Match 100.0%; Score 58; DB 19; Length 12;
 Best Local Similarity 100.0%; Pred. No. 0.00026;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 YRLAIRRLRLRY 12
 Db 1 YRLAIRRLRLRY 12
 RESULT 3
 AAM33799
 ID AAM33799 standard; peptide; 12 AA.
 AC AAM33799;
 XX
 DT 19-UN-1998 (first entry)
 DE Immunomodulating dimer peptide #3.
 XX
 KW Immunomodulating dimer; immunosuppressant drug; CTL activation;
 KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;
 KW rejection.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 PN MO9744351-A1.
 XX
 PD 27-NOV-1997.
 XX
 PF 22-MAY-1997; 97WO-US08689.
 XX
 PR 24-MAY-1996; 96US-0653294.
 XX
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 XX
 PI Beulow R, Clayberger C, Krensky AM;
 XX
 DR WPI; 1998-086530/08.
 XX
 PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B
 PT alpha-1 domain, used for preventing rejection of transplants or
 PT treating autoimmune diseases
 XX
 PS Claim 17; Page 35; 41pp; English.
 XX
 CC This sequence represents a specifically claimed immunomodulating
 CC dimer peptide of the invention. A peptide-type compound or variant is
 CC claimed which has immunomodulating activity, including the N-terminal
 CC acylated and/or C-terminal amidated or esterified forms of up to 60
 CC amino acids, where the peptide-type compound comprises the formula: A-B,
 CC where A, B = (R aa76-77L) (aa79-84) or (aa84-79) (aa77-76R); aa76 = E or

CC V; aa77 = D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a
 CC hydrophobic or small amino acid; aa82 = R or L; aa83 = G or R; and aa
 CC represents amino acid. The sequence in the brackets may optionally be
 CC absent or truncated at any peptide type bond within the brackets. The
 CC compounds comprise amino acid sequences related to a Class I HLA-B
 CC alpha1 domain (positions 79-84). They can be used to inhibit cytotoxic
 CC T-lymphocytes (CTL) from undesirably attacking cells in a host or in
 CC vitro. They can also be used in combination with antigenic peptides or
 CC proteins of interest to activate CTLs. They can also inhibit the
 CC proliferation of T cells in response to anti-CD3. The peptide can be
 CC used for preventing rejection of transplants or for treating autoimmune
 CC diseases, e.g. diabetes, rheumatoid arthritis and lupus erythematosus.
 CC The products can also be used for detection and diagnosis.

XX Sequence 12 AA;

Query Match 100.0%; Score 58; DB 19; Length 12;
 Best Local Similarity 100.0%; Pred. No. 0.00026;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLAIRRIALRY 12
 |||||
 Db 1 YRLAIRRIALRY 12

RESULT 4
 AAR95428
 ID AAR95428 standard; peptide; 20 AA.

XX AC AAR95428;

XX DT 12-NOV-1996 (first entry)

XX DE HLA-B2702 84-75-84 palindrome.

XX HLA; p74; alpha1-helix; human-leucocyte-associated antigen; inhibitor;
 KW T-cell lysate; membrane protein; mammal; heat shock protein; Hsc70; APC;
 KW B cell; calcium influx; cytotoxic T lymphocyte; CTL; differentiation;
 KW cytotoxicity; antigen presenting cell.

XX OS Synthetic.

XX PN WO9513288-Al.

XX PD 18-MAY-1995.

XX PF 10-NOV-1994; 94WO-US12985.

XX PR 10-NOV-1993; 93US-0150493.

XX (STRD) UNIV LELAND STANFORD JUNIOR.

XX Clayberger C, Krensky AM;

XX WPI; 1995-194027/25.

XX Compens. comprising lymphoid surface membrane proteins - which may
 PT inhibit cytolytic activity and differentiation of CTLs.

XX Example; Page 12; 29pp; English.

XX AAR95413, and AAR95415-R95431 represent palindromes and fragments of
 CC human-leucocyte-associated antigens. This sequence represents the
 CC HLA-B2702 84-75-84 palindrome. These sequences can be used to isolate
 CC the protein p74 from a T-cell lysate. p74 is a T-cell surface membrane
 CC protein associated with T-cell activation in mammalian T-cells, and is
 CC also immunologically cross reactive with the heat shock protein Hsc70.
 CC p74 is found in a limited number of cell types, but is particularly
 CC expressed on B and T cells. p74 can be isolated by lysis of a suitable
 CC cell with an amphoteric detergent, and then passed through an affinity
 CC column containing a covalently bound HLA-B2702 palindromic peptide.
 CC Compositions comprising the extracellular fragment of p74 combined with
 CC HLA-B2702.60-84 (see AAR95416), induces calcium influx, and inhibits

CC cytotoxic T lymphocyte (CTL) differentiation or cytotoxicity. Candidate
 CC compounds can be screened for their effect on the cytolytic activity of
 CC T-cells, by combining them with the extracellular portion of p74 and
 CC determining the amount of binding between the candidate compound and
 CC p74. Modulation of CTL activity can be inhibited in a cellular
 CC composition containing T-cells and antigen presenting cells (APCs), by
 CC adding to the mix the extracellular portion of p74, in an amount
 CC sufficient to compete with p74 for the binding of the p74 ligand.

XX Sequence 20 AA;

Query Match 75.9%; Score 44; DB 16; Length 20;
 Best Local Similarity 60.0%; Pred. No. 0.17;
 Matches 12; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 YRLAIR-----RIALRY 12
 |||||
 Db 1 YRLAIRLNERRRIALRY 20

RESULT 5

AAR92907

ID AAR92907 standard; peptide; 20 AA.

XX AC AAR92907;

XX DT 16-MAY-1996 (first entry)

XX DE HLA-B2702 CTL modulating peptide (B2702.84-75/75-84).

XX Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;
 KW immunosuppressant; graft versus host disorder; transplantation; therapy;
 KW class I MHC; HLA-B2702.

XX OS Synthetic.

XX PN WO9526979-Al.

XX PD 12-OCT-1995.

XX PF 05-APR-1995; 95WO-US04349.

XX PR 05-APR-1994; 94US-0222851.

XX (STRD) UNIV LELAND STANFORD JUNIOR.

XX Clayberger C, Krensky AM, Parham P;

XX WPI; 1995-358582/46.

XX Extension of acceptance period of transplants from MHC unmatched
 PT donor hosts - using Class I B75-84 MHC antigen of the recipient
 PT host

XX Example 15; Page 36; 80pp; English.

XX AAR83061-R83085, AAR83090-R83096 and AAR92907-R92914 represent fragments
 CC of class I major histocompatibility complex (MHC) antigens. This
 CC sequence is an inverted dimer of residues 75-84 of the alpha-1 domain of
 CC the class I MHC HLA-B2702. These sequences can be used to extend the
 CC period of acceptance by a recipient of a transplant from an MHC unmatched
 CC donor. The peptides are administered to a patient in conjunction with a
 CC subtherapeutic amount of an immunosuppressant. This is administered to
 CC the patient for a limited period of time (compared to the lifetime
 CC administration for current treatments). The peptides particularly
 CC modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)
 CC of the patient.

XX Sequence 20 AA;

Query Match 75.9%; Score 44; DB 16; Length 20;
 Best Local Similarity 60.0%; Pred. No. 0.17;
 Matches 12; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

OY 1 YRLAIR-----RIALRY 12
 |||||
 DB 1 YRLAIRLNERRRRLTALRY 20

RESULT 6

AAW33778
 ID AAW33778 standard; peptide, 20 AA.

AAW33778;

19-JUN-1998 (first entry)

Immunomodulating dimer peptide #1.

Immunomodulating dimer; immunosuppressant drug; CTL activation;
 transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;
 rejection.

Synthetic.

OS Homo sapiens.

MO9744351-A1.

27-NOV-1997.

22-MAY-1997; 97WO-US08689.

24-MAY-1996; 96US-0653294.

(STRD) UNIV LELAND STANFORD JUNIOR.

Beulow R, Clayberger C, Krensky AM;

WPI; 1998-086530/08.

New immunomodulating dimer peptide(s) - based on a Class I HLA-B
 alpha-1 domain, used for preventing rejection of transplants or
 treating autoimmune diseases

Claim 16; Page 35; 41pp; English.

This sequence represents a specifically claimed immunomodulating
 dimer peptide of the invention. A peptide-type compound or variant is
 claimed which has immunomodulating activity, including the N-terminal
 acylated and/or C-terminal amidated or esterified forms of up to 60
 amino acids, where the peptide-type compound comprises the formula: A-B,
 where A, B = (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or
 V; aa77 = D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a
 hydrophobic or small amino acid; aa82 = R or L; aa83 = G or R; and aa
 represents amino acid. The sequence in the brackets may optionally be
 absent or truncated at any peptide type bond within the brackets. The
 compounds comprise amino acid sequences related to a Class I HLA-B
 alpha1 domain (positions 79-84). They can be used to inhibit cytotoxic
 T-lymphocytes (CTL) from undesirably attacking cells in a host or in
 vitro. They can also be used in combination with antigenic peptides or
 proteins of interest to activate CTLs. They can also inhibit the
 proliferation of T cells in response to anti-CD3. The peptide can be
 used for preventing rejection of transplants or for treating autoimmune
 diseases, e.g. diabetes, rheumatoid arthritis and lupus erythematosus.
 The products can also be used for detection and diagnosis.

Sequence 20 AA;

Query Match 75.9%; Score 44; DB 19; Length 20;

Best Local Similarity 60.0%; Pred. No. 0.17; Mismatches 0; Indels 8; Gaps 1;

OY 1 YRLAIR-----RIALRY 12
 |||||
 DB 1 YRLAIRLNERRRRLTALRY 20

RESULT 7
 AAR92909
 ID AAR92909 standard; peptide, 20 AA.

AAR92909;

16-MAY-1996 (first entry)

HLA-B2702 CTL modulating peptide (B2702.84-75/75-84(T)).

Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;
 immunosuppressant; graft versus host disorder; transplantation; therapy;
 class I MHC; HLA-B2702.

Synthetic.

MO9526979-A1.

12-OCT-1995.

05-APR-1995; 95WO-US04349.

05-APR-1994; 94US-0222851.

(STRD) UNIV LELAND STANFORD JUNIOR.

Clayberger C, Krensky AM, Parham P;

WPI; 1995-358582/46.

Extension of acceptance period of transplants from MHC unmatched
 PT donor hosts - using Class I B75-84 MHC antigen of the recipient
 host

Example 15; Page 36; 80pp; English.

AA83061-R83085, AAR83090-R83096 and AAR92907-R92913 represent fragments
 of class I major histocompatibility complex (MHC) antigens. This
 sequence is an inverted dimer of residues 75-84 of the alpha-1 domain of
 the class I MHC HLA-B2702. These sequences can be used to extend the
 period of acceptance for a recipient of a transplant from an MHC unmatched
 donor. The peptides are administered to a patient in conjunction with a
 subtherapeutic amount of an immunosuppressant. This is administered to
 the patient for a limited period of time (compared to the lifetime
 administration for current treatments). The peptides particularly
 CC modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)
 of the patient.

Sequence 20 AA;

Query Match 67.2%; Score 39; DB 16; Length 20;

Best Local Similarity 55.0%; Pred. No. 1.4; Mismatches 1; Indels 8; Gaps 1;

OY 1 YRLAIR-----RIALRY 12
 |||||
 DB 1 YRLAIRLNERRRRLTALRY 20

RESULT 8

AAR92908
 ID AAR92908 standard; peptide, 20 AA.

AAR92908;

16-MAY-1996 (first entry)

HLA-B2702 CTL modulating peptide (B2702.84-75(T)/75-84).

Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;
 immunosuppressant; graft versus host disorder; transplantation; therapy;
 class I MHC; HLA-B2702.


```

PT alpha-1 domain, used for preventing rejection of transplants or
PT creating autoimmune diseases
XX
XX Example 1; Page 19; 41pp; English.
XX
CC Peptides AAW3784-98 and AAW3778-9 were assayed for their
CC immunomodulating activity. A peptide-type compound or variant is claimed
CC which has immunomodulating activity, including the N-terminal acylated
CC and/or C-terminal amidated or esterified forms of up to 60 amino acids,
CC where the peptide-type compound comprises the formula: A-B, where A, B =
CC (R aa76-77L) (aa79-84) or (aa84-79) (aa77-76R); aa76 = E or V; aa77 =
CC D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a hydrophobic or
CC small amino acid; aa82 = R or L; aa83 = G or R; and aa represents amino
CC acid. The sequence in the brackets may optionally be absent or truncated
CC at any peptide type bond within the brackets. The compounds comprise
CC amino acid sequences related to a Class I HLA-B alpha domain (positions
CC 79-84). They can be used to inhibit cytotoxic T-lymphocytes (CTL) from
CC undesirably attacking cells in a host or in vitro. They can also be
CC used in combination with antigenic peptides or proteins of interest to
CC activate CTLs. They can also inhibit the proliferation of T cells in
CC response to anti-CD3. The peptide can be used for preventing rejection
CC of transplants or for treating autoimmune diseases, e.g. diabetes,
CC rheumatoid arthritis and lupus erythematosus. The products can also be
CC used for detection and diagnosis.
CC
XX
SQ Sequence 20 AA;
Query Match 67.2%; Score 39; DB 19; Length 20;
Best Local Similarity 55.0%; Pred. No. 1.4;
Matches 11; Conservative 0; Mismatches 1; Indels 8; Gaps 1;
OY 1 YRLAIR-----RIALRY 12
1 YRLAIRLNERNRLRLRLRY 20
DB 1 YRLAIRLNERNRLRLRLRY 20
RESULT 11
AAR95430
ID AAR95430 standard; peptide; 20 AA.
XX
AC AAR95430;
XX
DT 12-NOV-1996 (first entry)
XX
DE HLA-B2702 84-75T/75-84T palindrome.
XX
KM HLA, p74; alpha-helix; human-leucocyte-associated antigen; inhibitor;
KM T-cell lysate; membrane protein; mammal; heat shock protein; Hsc70, ApC;
KM B cell; calcium influx; cytotoxic T lymphocyte; CTL; differentiation;
KM cytolytic; antigen presenting cell.
XX
OS Synthetic.
XX
PN WO9513288-A1.
XX
PD 18-MAY-1995.
XX
PF 10-NOV-1994; 94WO-US12985.
XX
PR 10-NOV-1993; 93US-0150493.
XX
PA (STRD ) UNIV LELAND STANFORD JUNIOR.
XX
PI Clayberger C, Krensky AM;
XX
DR WPI, 1995-194027/25.
XX
XX Compens. comprising lymphoid surface membrane proteins - which may
PT inhibit cytolytic activity and differentiation of CTLs.
XX
PS Example; Page 12; 29pp; English.
XX
CC AAR95413, and AAR95415-B95431 represent palindromes and fragments of

```

```

CC human-leucocyte-associated antigens. This sequence represents the
CC HLA-B2702 84-75T/75-84T palindrome. These sequences can be used to
CC isolate the protein p74 from a T-cell lysate. p74 is a T-cell surface
CC membrane protein associated with T-cell activation in mammalian T-cells,
CC and is also immunologically cross reactive with the heat shock protein
CC Hsc70. p74 is found in a limited number of cell types, but is
CC particularly expressed on B and T cells. p74 can be isolated by lysis
CC of a suitable cell with an amphoteric detergent, and then passed through
CC an affinity column containing a covalently bound HLA-B2702 palindromic
CC peptide. Compositions comprising the extracellular fragment of p74
CC combined with HLA-B2702.60-84 (see AAR95416), induces calcium influx,
CC and inhibits cytotoxic T lymphocyte (CTL) differentiation or cytotoxic
CC. Candidate compounds can be screened for their effect on the cytolytic
CC activity of T-cells, by combining them with the extracellular portion of
CC p74 and determining the amount of binding between the candidate compound
CC and p74. Modulation of CTL activity can be inhibited in a cellular
CC composition containing T-cells and antigen presenting cells (APCs), by
CC adding to the mix the extracellular portion of p74, in an amount
CC sufficient to compete with p74 for the binding of the p74 ligand.
CC
XX
SQ Sequence 20 AA;
Query Match 63.8%; Score 37; DB 16; Length 20;
Best Local Similarity 57.9%; Pred. No. 3.3;
Matches 11; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
OY 1 YRLAIR-----RIALRY 11
1 YRLAIRLNERNRLRLRLRY 19
DB 1 YRLAIRLNERNRLRLRLRY 19
RESULT 12
AAR71429
ID AAR71429 standard; peptide; 18 AA.
XX
AC AAR71429;
XX
DT 12-OCT-1995 (first entry)
XX
DE Human MHC 1 alpha 1 domain peptide [Ala81]-Dk-(69-85).
XX
KM Major histocompatibility complex class 1; MHC 1; cell receptor;
KM alpha 1 domain; peptide [Ala81]-Dk-(69-85); interaction modulation;
KM arthritis; neoplasias; lupus erythematosus.
XX
OS Homo sapiens.
XX
PN WO9505189-A.
XX
PD 23-FEB-1995.
XX
PF 12-AUG-1994; 94WO-US09189.
XX
PR 12-AUG-1993; 93US-0105416.
XX
PA (REGC ) UNIV CALIFORNIA.
XX
PI Goldstein A, Goodenow RS, Olsson L;
XX
DR WPI, 1995-098577/13.
XX
XX Regulating cell surface receptor response - by modulating
PT interaction between MHC class I antigen and the cell surface
PT receptor
XX
PS Example 4; Page 45; 103pp; English.
XX
XX AAR71424-R71438 are human major histocompatibility complex class 1
CC (MHC 1) alpha 1 domain derived peptides and peptide mutants, they
CC were used to modulate interactions between MHC 1 and cell surface
CC receptors. Via competitive inhibition the peptides diminish the
CC receptors response, this feature may be useful for the treatment
CC of neoplasias, lupus erythematosus and arthritis.

```


PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 PR XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 PI XX
 DR MPI; 2001-483426/52.
 DR N-PSDB; AAK55864.
 XX
 PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and
 PT metastasis -
 PT XX
 PS Claim 11; SEQ ID NO 10676; 3071bp + Sequence Listing; English.
 PS XX
 CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
 CC amino acid sequences given in AAM62170 to AAM91921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patient's own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting
 CC the nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/hematopoietic-related diseases, especially
 CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/hematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAM62169
 CC represent sequences used in the exemplification of the present invention.
 CC XX
 SQ Sequence 46 AA;
 Query Match 60.3%; Score 35; DB 22; Length 46;
 Best Local Similarity 60.0%; Pred. No. 18;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 OY 1 YRLAIRRIAL 10
 Db 9 YRLGVRMAL 18
 RESULT 14
 AAR92911
 ID AAR92911 standard; peptide; 20 AA.
 XX
 AC AAR92911;
 XX
 DT 16-MAY-1996 (first entry)

XX HLA-B2702 CTL modulating peptide (B2702.84-75/84-75).
 DE Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;
 XX immunopressant; graft versus host disorder; transplantation; therapy;
 KW class I MHC; HLA-B2702.
 KW
 XX Synthetic.
 OS WO9526979-A1.
 XX
 PN 12-OCT-1995.
 PD
 XX 05-APR-1995; 95WO-US04349.
 PF
 XX 05-APR-1994; 94US-0222851.
 PR
 XX (STRD) UNIV LELAND STANFORD JUNIOR.
 PA Clayberger C, Krensky AM, Parham P;
 PI
 XX WPI; 1995-358582/46.
 DR
 XX
 PT Extension of acceptance period of transplants from MHC unmatched
 PT donor hosts - using Class I B75-84 MHC antigen of the recipient
 PT host
 PT XX
 PS Example 15; Page 36; 80pp; English.
 PS XX
 CC AAR83061-R83085, AAR83090-R83096 and AAR92907-R92914 represent fragments
 CC of class I major histocompatibility complex (MHC) antigens. This
 CC sequence is a dimer of residues 75-84 of the alpha-1 domain of the class
 CC I MHC HLA-B2702. These sequences can be used to extend the period of
 CC acceptance by a recipient of a transplant from an MHC unmatched donor.
 CC The peptides are administered to a patient in conjunction with a
 CC subtherapeutic amount of an immunosuppressant. This is administered to
 CC the patient for a limited period of time (compared to the lifetime
 CC administration for current treatments). The peptides particularly
 CC modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)
 CC of the patient.
 CC XX
 SQ Sequence 20 AA;
 Query Match 59.5%; Score 34.5; DB 16; Length 20;
 Best Local Similarity 56.2%; Pred. No. 9.7;
 Matches 9; Conservative 2; Mismatches 0; Indels 5; Gaps 1;
 OY 1 YRLAIR-----RIAIR 11
 Db 1 YRLAIRLNERIRLAIIR 16
 RESULT 15
 AAM33779
 ID AAM33779 standard; peptide; 20 AA.
 XX
 AC AAM33779;
 XX
 DT 19-JUN-1998 (first entry)
 XX
 DE Immunomodulating dimer peptide #2.
 XX
 KW Immunomodulating dimer; immunosuppressant drug; CTL activation;
 KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;
 KW rejection.
 KW
 XX Synthetic.
 OS Homo sapiens.
 OS WO9744351-A1.
 PN
 XX
 PD 27-NOV-1997.
 XX

PF 22-MAY-1997; 97WO-US08689.
 XX
 PR 24-MAY-1996; 96US-0653294.
 XX
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 XX
 PI Beulow R, Clayberger C, Krensky AM;
 XX WPI; 1998-086530/08.
 DR
 XX New immunomodulating dimer peptide(s) - based on a Class I HLA-B
 PT alpha-1 domain, used for preventing rejection of transplants or
 PT treating autoimmune diseases
 XX
 PS Claim 16; Page 35; 41pp; English.
 XX
 CC This sequence represents a specifically claimed immunomodulating
 CC dimer peptide of the invention. A peptide-type compound or variant is
 CC claimed which has immunomodulating activity, including the N-terminal
 CC acylated and/or C-terminal amidated or esterified forms of up to 60
 CC amino acids, where the peptide-type compound comprises the formula; A-B,
 CC where A, B = (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or
 CC V; aa77 = D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a
 CC hydrophobic or small amino acid; aa82 = R or L; aa83 = G or R; and aa
 CC represents amino acid. The sequence in the brackets may optionally be
 CC absent or truncated at any peptide type bond within the brackets. The
 CC compounds comprise amino acid sequences related to a Class I HLA-B
 CC alphas domain (positions 79-84). They can be used to inhibit cytotoxic
 CC T-lymphocytes (CTL) from undesirably attacking cells in a host or in
 CC vitro. They can also be used in combination with antigenic peptides or
 CC proteins of interest to activate CTLs. They can also inhibit the
 CC proliferation of T cells in response to anti-CD3. The peptide can be
 CC used for preventing rejection of transplants or for treating autoimmune
 CC diseases, e.g. diabetes, rheumatoid arthritis and lupus erythematosus.
 CC The products can also be used for detection and diagnosis.
 XX
 SQ Sequence 20 AA;

Query Match 59.5%; Score 34.5; DB 19; Length 20;
 Best Local Similarity 56.2%; Pred. No. 9.7;
 Matches 9; Conservative 2; Mismatches 0; Indels 5; Gaps 1;
 Qy 1 YRLAIR-----RIALR 11
 ||||| :|||
 Db 1 YRLAIRLNERYRLAIR 16

Search completed: December 19, 2002, 16:19:05
 Job time : 29.5882 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 19, 2002, 16:16:27 ; Search time 9.70588 Seconds
(without alignments)
36.377 Million cell updates/sec

Title: US-08-653-294C-36
Perfect score: 58
Sequence: 1 YRLAIRRIALRY 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 185840

Minimum DB seq length: 0
Maximum DB seq length: 60

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/prodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/prodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/prodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/prodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/prodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	* Match	Length	ID	Description
1	58	100.0	12	4	US-08-651-650-28
2	58	100.0	12	5	PCT-US94-12985-5
3	44	75.9	20	1	US-08-222-851-33
4	44	75.9	20	4	US-08-651-650-18
5	44	75.9	20	5	PCT-US94-12985-1
6	44	75.9	20	5	PCT-US94-12985-4
7	39	67.2	20	1	US-08-222-851-34
8	39	67.2	20	1	US-08-222-851-35
9	39	67.2	20	4	US-08-651-650-20
10	39	67.2	20	4	US-08-651-650-21
11	37	63.8	20	5	PCT-US94-12985-6
12	35	60.3	18	1	US-08-105-416-16
13	35	60.3	18	2	US-08-473-656A-16
14	35	60.3	18	3	US-08-483-931B-16
15	34.5	59.5	20	1	US-08-222-851-37
16	34.5	59.5	20	4	US-08-651-650-23
17	34	58.6	20	1	US-08-222-851-36
18	34	58.6	20	4	US-08-651-650-22
19	32	55.2	40	2	US-08-679-405-6
20	32	55.2	40	2	US-08-842-799-6
21	32	55.2	40	5	PCT-US96-11458-6
22	30	51.7	52	1	US-08-346-849-8
23	30	51.7	52	2	US-08-293-284A-8
24	30	51.7	54	1	US-08-346-849-13
25	30	51.7	54	2	US-08-293-284A-13
26	29	50.0	6	4	US-08-651-650-30
27	29	50.0	6	4	US-08-651-650-32

28	29	50.0	9	4	US-08-433-613-50	Sequence 50, Appl
29	29	50.0	10	1	US-08-222-851-13	Sequence 13, Appl
30	29	50.0	10	1	US-08-222-851-30	Sequence 30, Appl
31	29	50.0	10	1	US-08-440-504A-4	Sequence 4, Appl
32	29	50.0	10	1	US-08-440-504A-5	Sequence 5, Appl
33	29	50.0	10	4	US-08-433-613-57	Sequence 57, Appl
34	29	50.0	10	4	US-08-651-650-12	Sequence 12, Appl
35	29	50.0	10	4	US-08-651-650-15	Sequence 15, Appl
36	29	50.0	10	4	US-08-651-650-34	Sequence 34, Appl
37	29	50.0	10	4	US-08-651-650-36	Sequence 36, Appl
38	29	50.0	10	5	PCT-US94-12985-9	Sequence 9, Appl
39	29	50.0	10	5	PCT-US94-12985-12	Sequence 12, Appl
40	29	50.0	11	4	US-08-433-613-26	Sequence 26, Appl
41	29	50.0	12	4	US-08-433-613-25	Sequence 25, Appl
42	29	50.0	13	4	US-08-433-613-24	Sequence 24, Appl
43	29	50.0	14	4	US-08-433-613-4	Sequence 4, Appl
44	29	50.0	14	4	US-08-433-613-7	Sequence 7, Appl
45	29	50.0	14	4	US-08-433-613-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1
US-08-651-650-28
; Sequence 28, Application US/08651650
; Patent No. 6436903
; GENERAL INFORMATION:
; APPLICANT: CLAYBERGER, Carol A.
; APPLICANT: KRENSKY, Alan M.
; TITLE OF INVENTION: IMMUNOMODULATING COMPOUNDS COMPRISING
; D-ISOMERS OF AMINO ACIDS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSER: MORRISON & FOERSTER
; STREET: 2000 PENNSYLVANIA AVENUE, NW
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1888
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/651,650
; FILING DATE: 22-MAY-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 28600-20203.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 822-0168
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; TOPOLOGY: single
; STRANDEDNESS: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: B2702.84-79/79-84
; US-08-651-650-28

Query Match 100.0%; Score 58; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 6.5e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLAIRRIALRY 12

Db 1 YRLAIRRLRY 12

RESULT 2
PCT-US94-12985-5
Sequence 5, Application PC/TUS9412985
GENERAL INFORMATION:
APPLICANT: The Board of Trustees for the Leland Stanford Junior
APPLICANT: University
TITLE OF INVENTION: SURFACE MEMBRANE PROTEINS AND THEIR
TITLE OF INVENTION: EFFECT ON IMMUNE RESPONSE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HORNBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/12985
FILING DATE: 10-NOV-1994
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/150,493
FILING DATE: 10-NOV-1993
ATTORNEY/AGENT INFORMATION:
NAME: Rowland, Bertam I
REGISTRATION NUMBER: 20,015
REFERENCE/DOCKET NUMBER: FP-58976-PC/BIR
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US94-12985-5

Query Match 100.0%; Score 58; DB 5; Length 12;
Best Local Similarity 100.0%; Pred. No. 6.5e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLAIRRLRY 12
Db 1 YRLAIRRLRY 12

RESULT 3
US-08-222-851-33
Sequence 33, Application US/08222851
Patent No. 5723128
GENERAL INFORMATION:
APPLICANT: CLAYBERGER, CAROL A.
APPLICANT: KRENSKY, ALAN M.
APPLICANT: PARHAM, PETER
TITLE OF INVENTION: CYTOTOXIC T-CELL LYMPHOCYTE ("CTL")
TITLE OF INVENTION: ACTIVITY REGULATION BY CLASS I MHC PEPTIDES
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 PENNSYLVANIA AVENUE, NW, STE 5500
CITY: WASHINGTON

STATE: DC
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/222,851
FILING DATE: 05-APR-1994
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: MILLMAN, ROBERT A.
REGISTRATION NUMBER: 36,217
REFERENCE/DOCKET NUMBER: 28600-20200.22
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 494-0792
TELEX: 90-4030 MRSNFOERSMSH
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-222-851-33

Query Match 75.9%; Score 44; DB 1; Length 20;
Best Local Similarity 60.0%; Pred. No. 0.043;
Matches 12; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 YRLAIR-----RLRY 12
Db 1 YRLAIRLNRRLRY 20

RESULT 4
US-08-651-650-18
Sequence 18, Application US/08651650
Patent No. 6436903
GENERAL INFORMATION:
APPLICANT: CLAYBERGER, CAROL A.
APPLICANT: KRENSKY, ALAN M.
TITLE OF INVENTION: IMMUNOMODULATING COMPOUNDS COMPRISING
TITLE OF INVENTION: D-ISOMERS OF AMINO ACIDS
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 PENNSYLVANIA AVENUE, NW
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20006-1888
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/651,650
FILING DATE: 22-MAY-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: MURASHIGE, KATE H.
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 28600-20203.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 822-0168
TELEX: 90-4030
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:

; LENGTH: 20 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; ORIGINAL SOURCE:
 ; INDIVIDUAL ISOLATE: B2702.84-75/75-84
 US-08-651-650-18

Query Match 75.9%; Score 44; DB 4; Length 20;
 Best Local Similarity 60.0%; Pred. No. 0.043;
 Matches 12; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

Qy 1 YRLAIR-----RIALRY 12
 |||||
 Db 1 YRLAIRLNERRENRIALRY 20

RESULT 5

PCT-US94-12985-1
 ; Sequence 1, Application PC/TUS9412985

; GENERAL INFORMATION:
 ; APPLICANT: The Board of Trustees for the Leland Stanford Junior
 ; APPLICANT: University
 ; TITLE OF INVENTION: SURFACE MEMBRANE PROTEINS AND THEIR
 ; TITLE OF INVENTION: EFFECT ON IMMUNE RESPONSE
 ; NUMBER OF SEQUENCES: 20
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
 ; STREET: 4 Embarcadero Center, Suite 3400
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94111-4187

; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US94/12985
 ; FILING DATE: 10-NOV-1994
 ; CLASSIFICATION:

; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/150,493
 ; FILING DATE: 10-NOV-1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Rowland, Bertram I
 ; REGISTRATION NUMBER: 20,015
 ; REFERENCE/DOCKET NUMBER: FP-58976-PC/BIR
 ; TELEPHONE: (415) 781-1989
 ; TELEFAX: (415) 398-3249
 ; TELEX: 910 277299
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 20 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 PCT-US94-12985-1

Query Match 75.9%; Score 44; DB 5; Length 20;
 Best Local Similarity 60.0%; Pred. No. 0.043;
 Matches 12; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

Qy 1 YRLAIR-----RIALRY 12
 |||||
 Db 1 YRLAIRLNERRENRIALRY 20

RESULT 6

PCT-US94-12985-4
 ; Sequence 4, Application PC/TUS9412985
 ; GENERAL INFORMATION:
 ; APPLICANT: The Board of Trustees for the Leland Stanford Junior
 ; APPLICANT: University
 ; TITLE OF INVENTION: SURFACE MEMBRANE PROTEINS AND THEIR
 ; TITLE OF INVENTION: EFFECT ON IMMUNE RESPONSE
 ; NUMBER OF SEQUENCES: 20
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
 ; STREET: 4 Embarcadero Center, Suite 3400
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94111-4187

; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US94/12985
 ; FILING DATE: 10-NOV-1994
 ; CLASSIFICATION:

; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/150,493
 ; FILING DATE: 10-NOV-1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Rowland, Bertram I
 ; REGISTRATION NUMBER: 20,015
 ; REFERENCE/DOCKET NUMBER: FP-58976-PC/BIR
 ; TELEPHONE: (415) 781-1989
 ; TELEFAX: (415) 398-3249
 ; TELEX: 910 277299
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 20 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 PCT-US94-12985-4

Query Match 75.9%; Score 44; DB 5; Length 20;
 Best Local Similarity 60.0%; Pred. No. 0.043;
 Matches 12; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

Qy 1 YRLAIR-----RIALRY 12
 |||||
 Db 1 YRLAIRLNERRENRIALRY 20

RESULT 7

US-08-222-851-34
 ; Sequence 34, Application US/08222851
 ; Patent No. 5723128
 ; GENERAL INFORMATION:
 ; APPLICANT: CLAYBERGER, CAROL A.
 ; APPLICANT: KRENSKY, ALAN N.
 ; APPLICANT: PARHAM, PETER
 ; TITLE OF INVENTION: CYTOTOXIC T-CELL LYMPHOCYTE ("CTL")
 ; TITLE OF INVENTION: ACTIVITY REGULATION BY CLASS I MHC PEPTIDES
 ; NUMBER OF SEQUENCES: 43
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: MORRISON & FOERSTER
 ; STREET: 2000 PENNSYLVANIA AVENUE, NW, STE 5500
 ; CITY: WASHINGTON
 ; STATE: DC
 ; COUNTRY: USA
 ; ZIP: 20006-1812
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/222,851
FILING DATE: 05-APR-1994
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: MILLMAN, ROBERT A.
REGISTRATION NUMBER: 36,217
REFERENCE/DOCKET NUMBER: 28600-20200.22
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 494-0792
TELEX: 90-4030 MRSNFOERSM
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-222-851-34

Query Match 67.2%; Score 39; DB 1; Length 20;
Best Local Similarity 55.0%; Pred. No. 0.36;
Matches 11; Conservative 0; Mismatches 1; Indels 8; Gaps 1;

QY 1 YRLAIR-----RIALRY 12
Db 1 YRLAIRLNRRERNTALRY 20

RESULT 8
US-08-222-851-35
Sequence 35, Application US/08222851
Patent No. 5723128
GENERAL INFORMATION:
APPLICANT: CLAYBERGER, CAROL A.
APPLICANT: KRENSKY, ALAN M.
APPLICANT: PARHAM, PETER
TITLE OF INVENTION: CYTOTOXIC T-CELL LYMPHOCYTE ("CTL")
TITLE OF INVENTION: ACTIVITY REGULATION BY CLASS I MHC PEPTIDES
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 PENNSYLVANIA AVENUE, NW, STE 5500
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/222,851
FILING DATE: 05-APR-1994
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: MILLMAN, ROBERT A.
REGISTRATION NUMBER: 36,217
REFERENCE/DOCKET NUMBER: 28600-20200.22
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 494-0792
TELEX: 90-4030 MRSNFOERSM
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

US-08-222-851-35

Query Match 67.2%; Score 39; DB 1; Length 20;
Best Local Similarity 55.0%; Pred. No. 0.36;
Matches 11; Conservative 0; Mismatches 1; Indels 8; Gaps 1;

QY 1 YRLAIR-----RIALRY 12
Db 1 YRLAIRLNRRERNTALRY 20

RESULT 9
US-08-651-650-20
Sequence 20, Application US/08651650
Patent No. 6436903
GENERAL INFORMATION:
APPLICANT: CLAYBERGER, Carol A.
APPLICANT: KRENSKY, Alan M.
TITLE OF INVENTION: IMMUNOMODULATING COMPOUNDS COMPRISING
TITLE OF INVENTION: D-ISOMERS OF AMINO ACIDS
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 PENNSYLVANIA AVENUE, NW
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20006-1888
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/651,650
FILING DATE: 22-MAY-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: MURASHIGE, KATE H.
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 28600-20203.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 822-0168
TELEX: 90-4030
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: B2702.84-75T/75-84
US-08-651-650-20

Query Match 67.2%; Score 39; DB 4; Length 20;
Best Local Similarity 55.0%; Pred. No. 0.36;
Matches 11; Conservative 0; Mismatches 1; Indels 8; Gaps 1;

QY 1 YRLAIR-----RIALRY 12
Db 1 YRLAIRLNRRERNTALRY 20

RESULT 10
US-08-651-650-21
Sequence 21, Application US/08651650
Patent No. 6436903
GENERAL INFORMATION:
APPLICANT: CLAYBERGER, Carol A.
APPLICANT: KRENSKY, Alan M.
TITLE OF INVENTION: IMMUNOMODULATING COMPOUNDS COMPRISING

;; TITLE OF INVENTION: D-ISOMERS OF AMINO ACIDS
;; NUMBER OF SEQUENCES: 38
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: MORRISON & FOERSTER
;; STREET: 2000 PENNSYLVANIA AVENUE, NW
;; CITY: WASHINGTON
;; STATE: DC
;; COUNTRY: USA
;; ZIP: 20006-1888
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; OPERATING SYSTEM: IBM PC compatible
;; SOFTWARE: Patent in Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; FILING DATE: 22-MAY-1996
;; APPLICATION NUMBER: US/08/651,650
;; CLASSIFICATION: 536
;; ATTORNEY/AGENT INFORMATION:
;; NAME: MURASHIGE, KATE H.
;; REGISTRATION NUMBER: 29,959
;; REFERENCE/DOCKET NUMBER: 28600-20203.00
;; TELEPHONE: (202) 897-1500
;; TELEFAX: (202) 822-0168
;; TELEX: 90-4030
;; INFORMATION FOR SEQ ID NO: 21:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 20 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; ORIGINAL SOURCE:
;; INDIVIDUAL ISOLATE: B7202.84-75/75-84T
US-08-651-650-21

Query Match 67.2%; Score 39; DB 4; Length 20;
Best Local Similarity 55.0%; Pred. No. 0.36;
Matches 11; Conservative 0; Mismatches 1; Indels 8; Gaps 1;

QY 1 YRLAIR-----RIALRY 12
||| ||| ||| |||
Db 1 YRLAIRLNRERLRLTALRY 20

RESULT 11
PCT-US94-12985-6
; Sequence 6, Application PC/TUS9412985
; GENERAL INFORMATION:
; APPLICANT: The Board of Trustees for the Leland Stanford Junior
; APPLICANT: University
; TITLE OF INVENTION: SURFACE MEMBRANE PROTEINS AND THEIR
; TITLE OF INVENTION: EFFECT ON IMMUNE RESPONSE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/12985
; FILING DATE: 10-NOV-1994
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/150,493

;; FILING DATE: 10-NOV-1993
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Rowland, Bertram I.
;; REGISTRATION NUMBER: 20,015
;; REFERENCE/DOCKET NUMBER: FP-58976-PC/BIR
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 781-1989
;; TELEFAX: (415) 398-3249
;; TELEX: 910 277299
;; INFORMATION FOR SEQ ID NO: 6:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 20 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
PCT-US94-12985-6

Query Match 63.8%; Score 37; DB 5; Length 20;
Best Local Similarity 57.9%; Pred. No. 0.84;
Matches 11; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 YRLAIR-----RIALR 11
||| ||| ||| |||
Db 1 YRLAIRLNRERLRLTALR 19

RESULT 12
US-08-105-416-16
; Sequence 16, Application US/08105416
; Patent No. 5639958
; GENERAL INFORMATION:
; APPLICANT: Oleson, Lennart
; APPLICANT: Goodenow, Robert S
; APPLICANT: Goldstein, Avram
; TITLE OF INVENTION: Class I MHC Modulation of Surface
; TITLE OF INVENTION: Receptor Activity
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: US
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/105,416
; FILING DATE: 12-AUG-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Rowland, Bertram I.
; REGISTRATION NUMBER: 20015
; REFERENCE/DOCKET NUMBER: A55115-4/BIR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-781-1989
; TELEFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-105-416-16

Query Match 60.3%; Score 35; DB 1; Length 18;
Best Local Similarity 50.0%; Pred. No. 1.7;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/222,851
FILING DATE: 05-APR-1994
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: MILLMAN, ROBERT A.
REGISTRATION NUMBER: 36,217
REFERENCE/DOCKET NUMBER: 28600-20200.22
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 494-0792
TELEX: 90-4030 MRSNFOERSWSH
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-222-851-37

Query Match 59.5%; Score 34.5; DB 1; Length 20;
Best Local Similarity 56.2%; Pred. No. 2.4;
Matches 9; Conservative 2; Mismatches 0; Indels 5; Gaps 1;

QY 1 YRLAIR-----RIALR 11
|||
Db 1 YRLAIRLNERYRLAIR 16

Search completed: December 19, 2002, 16:23:57
Job time : 9.70588 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 19, 2002, 16:02:26 ; Search time 28.5982 Seconds
(without alignments)
55.932 Million cell updates/sec

Title: US-08-653-294C-42

Perfect score: 58

Sequence: 1 YRLIRYRLAIR 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 473727

Minimum DB seq length: 0

Maximum DB seq length: 60

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_101002.*
1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	41	70.7	20	16 AAR92911	HLA-B2702 CTL modu
2	41	70.7	20	19 AAW33779	Immunomodulating d
3	31.5	54.3	12	16 AAR95429	HLA-B2702 84-75-84
4	31.5	54.3	12	19 AAW33798	Peptide B2702.84-7
5	31.5	54.3	12	19 AAW33799	Immunomodulating d
6	31	53.4	28	22 ABB29624	Peptide #2275 enco
7	31	53.4	28	22 ABB34801	Peptide #2307 enco
8	31	53.4	28	22 ABB20217	Protein #2216 enco
9	31	53.4	28	22 AAW55603	Human brain expres
10	31	53.4	28	22 AAM67990	Human bone marrow

11	31	53.4	28	22 AAM15805	Peptide #2239 enco
12	31	53.4	28	22 AAM28314	Peptide #2351 enco
13	31	53.4	28	22 AAM03543	Peptide #2225 enco
14	31	53.4	28	23 ABG37525	Human polypeptide
15	31	53.4	50	22 AAO01990	Human peptide #428
16	30	51.7	20	22 ABB27777	Peptide #454 enco
17	30	51.7	20	22 ABB32948	Protein #421 enco
18	30	51.7	20	22 ABB18422	Human bone marrow
19	30	51.7	20	22 AAM53749	Human brain expres
20	30	51.7	20	22 AAM66132	Human bone marrow
21	30	51.7	20	22 AAM14003	Peptide #437 enco
22	30	51.7	20	22 AAM26409	Peptide #446 enco
23	30	51.7	20	22 AAM01744	Peptide #426 enco
24	30	51.7	20	23 ABG35782	Human peptide enco
25	30	51.7	37	22 ABB39053	Human brain expres
26	30	51.7	37	22 AAM59705	Human brain expres
27	30	51.7	37	22 AAM72285	Human bone marrow
28	30	51.7	37	22 AAM32544	Peptide #6581 enco
29	30	51.7	37	23 ABG42100	Human peptide enco
30	30	51.7	55	22 AAU39568	Propionibacterium
31	30	51.7	56	22 AAM87587	Human immune/haema
32	29	50.0	6	19 AAW33783	Peptide #4 used in
33	29	50.0	6	19 AAW33781	Peptide #2 used in
34	29	50.0	6	19 AAW47263	Immunomodulatory p
35	29	50.0	6	19 AAW47264	Immunomodulatory p
36	29	50.0	10	19 AAW82826	Cytomodulating lip
37	29	50.0	10	19 AAW47266	Immunomodulatory p
38	29	50.0	10	19 AAW47268	Immunomodulatory p
39	29	50.0	10	19 AAW47270	Immunomodulatory p
40	29	50.0	10	19 AAW47272	Immunomodulatory p
41	29	50.0	10	22 AAY72478	Immunosuppressive
42	29	50.0	10	22 AAY72481	Immunosuppressive
43	29	50.0	11	19 AAW76012	LM609 grafted anti
44	29	50.0	11	22 AAB61370	LM609 VL CDR2 pept
45	29	50.0	20	16 AAR95428	HLA-B2702 84-75-84

ALIGNMENTS

RESULT 1

AAR92911
ID AAR92911 standard; peptide; 20 AA.

XX AAR92911;

AC AAR92911;

DT 16-MAY-1996 (first entry)

XX HLA-B2702 CTL modulating peptide (B2702.84-75/84-75).

XX Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;

KW immunosuppressant; graft versus host disorder; transplantation; therapy;

KW class I MHC; HLA-B2702.

XX Synthetic.

OS WO9526979-A1.

XX PN 12-OCT-1995.

XX PD 05-APR-1995; 95WO-004349.

XX PF 05-APR-1994; 94US-0222851.

XX PR (STRD) UNIV LELAND STANFORD JUNIOR.

XX PA Clayberger C, Krensky AM, Parham P;

XX PI WPI; 1995-358582/46.

XX DR Extension of acceptance period of transplants from MHC unmatched

PT donor hosts - using Class I B75-84 MHC antigen of the recipient

PT host

XX Example 15; Page 36; 80pp; English.
 PS
 XX AAR83061-R83065, AAR83090-R83096 and AAR92907-R92914 represent fragments
 CC of class I major histocompatibility complex (MHC) antigens. This
 CC sequence is a dimer of residues 75-84 of the alpha-1 domain of the class
 CC I MHC HLA-B*2702. These sequences can be used to extend the period of
 CC acceptance by a recipient of a transplant from an MHC unmatched donor.
 CC The peptides are administered to a patient in conjunction with a
 CC subtherapeutic amount of an immunosuppressant. This is administered to
 CC the patient for a limited period of time (compared to the lifetime
 CC administration for current treatments). The peptides particularly
 CC modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)
 CC of the patient.
 CC
 SQ Sequence 20 AA;
 XX
 Query Match 70.7%; Score 41; DB 16; Length 20;
 Best Local Similarity 68.8%; Pred. No. 0.38;
 Matches 11; Conservative 0; Mismatches 1; Indels 4; Gaps 1;
 QY 1 YRLLI----RYRLAIR 12
 ||| ||| ||| ||| |||
 Db 1 YRLAIRLNERVRLAIR 16
 RESULT 2
 AAM33779
 ID AAM33779 standard; peptide; 20 AA.
 XX
 AC AAM33779;
 XX
 DT 19-JUN-1998 (first entry)
 XX
 DE Immunomodulating dimer peptide #2.
 XX
 KM Immunomodulating dimer; immunosuppressant drug; CTL activation;
 KM transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;
 KM rejection.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 PN WO9444351-A1.
 XX
 PD 27-NOV-1997.
 XX
 PF 22-MAY-1997; 97WO-US08689.
 XX
 PR 24-MAY-1996; 96US-0653294.
 XX
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 XX
 PI Beulow R, Clayberger C, Krensky AM;
 XX
 DR WPI; 1998-086530/08.
 XX
 PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B
 PT alpha-1 domain, used for preventing rejection of transplants or
 PT treating autoimmune diseases
 XX
 PS Claim 16; Page 35; 41pp; English.
 CC This sequence represents a specifically claimed immunomodulating
 CC dimer peptide of the invention. A peptide-type compound or variant is
 CC claimed which has immunomodulating activity, including the N-terminal
 CC acylated and/or C-terminal amidated or esterified forms of up to 60
 CC amino acids, where the peptide-type compound comprises the formula; A-B,
 CC where A, B = (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or
 CC V; aa77 = D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a
 CC hydrophobic or small amino acid; aa82 = R or L; aa83 = G or R; and aa
 CC represents amino acid. The sequence in the brackets may optionally be
 CC absent or truncated at any peptide type bond within the brackets. The

CC compounds comprise amino acid sequences related to a Class I HLA-B
 CC alpha domain (positions 79-84). They can be used to inhibit cytotoxic
 CC T-lymphocytes (CTL) from undesirably attacking cells in a host or in
 CC vitro. They can also be used in combination with antigenic peptides or
 CC proteins of interest to activate CTLs. They can also inhibit the
 CC proliferation of T cells in response to anti-CD3. The peptide can be
 CC used for preventing rejection of transplants or for treating autoimmune
 CC diseases, e.g. diabetes, rheumatoid arthritis and lupus erythematosus.
 CC The products can also be used for detection and diagnosis.
 CC
 SQ Sequence 20 AA;
 XX
 Query Match 70.7%; Score 41; DB 19; Length 20;
 Best Local Similarity 68.8%; Pred. No. 0.38;
 Matches 11; Conservative 0; Mismatches 1; Indels 4; Gaps 1;
 QY 1 YRLLI----RYRLAIR 12
 ||| ||| ||| ||| |||
 Db 1 YRLAIRLNERVRLAIR 16
 RESULT 3
 AAR95429
 ID AAR95429 standard; peptide; 12 AA.
 XX
 AC AAR95429;
 XX
 DT 12-NOV-1996 (first entry)
 XX
 DE HLA-B*2702 84-79-84 palindrome.
 XX
 KM HLA, p74; alpha-helix; human-leucocyte-associated antigen; inhibitor;
 KM T-cell lysate; membrane protein; mammal; heat shock protein; Hsc70; APC;
 KM B cell; calcium influx; cytotoxic T lymphocyte; CTL; differentiation;
 KM cycloysis; antigen presenting cell.
 XX
 OS Synthetic.
 OS WO9513288-A1.
 XX
 PD 18-MAY-1995.
 XX
 PF 10-NOV-1994; 94WO-US12985.
 XX
 PR 10-NOV-1993; 93US-0150493.
 XX
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 XX
 PI Clayberger C, Krensky AM;
 XX
 DR WPI; 1995-194027/25.
 XX
 PT Compns. comprising lymphoid surface membrane proteins - which may
 PT inhibit cytolytic activity and differentiation of CTLs.
 XX
 PS Example; Page 12; 29pp; English.
 CC AAR95413, and AAR95415-R95431 represent palindromes and fragments of
 CC human-leucocyte-associated antigens. This sequence represents the
 CC HLA-B*2702 84-79-84 palindrome. These sequences can be used to isolate
 CC the protein p74 from a T-cell lysate. p74 is a T-cell surface membrane
 CC protein associated with T-cell activation in mammalian T-cells, and is
 CC also immunologically cross reactive with the heat shock protein Hsc70.
 CC p74 is found in a limited number of cell types, but is particularly
 CC expressed on B and T cells. p74 can be isolated by lysis of a suitable
 CC cell with an amphoteric detergent, and then passed through an affinity
 CC column containing a covalently bound HLA-B*2702 palindromic peptide.
 CC Compositions comprising the extracellular fragment of p74 combined with
 CC HLA-B*2702-60-84 (see AAR95416), induces calcium influx, and inhibits
 CC cytotoxic T lymphocyte (CTL) differentiation or cycloysis. Candidate
 CC compounds can be screened for their effect on the cycloytic activity of
 CC T-cells, by combining them with the extracellular portion of p74 and
 CC determining the amount of binding between the candidate compound and

CC p74.. Modulation of CTL activity can be inhibited in a cellular
 CC composition containing T-cells and antigen presenting cells (APCs), by
 CC adding to the mix the extracellular portion of p74, in an amount
 CC sufficient to compete with p74 for the binding of the p74 ligand.
 XX
 SQ Sequence 12 AA;
 Query Match 54.3%; Score 31.5; DB 16; Length 12;
 Best Local Similarity 66.7%; Pred. No. 14;
 Matches 8; Conservative 2; Mismatches 1; Indels 1; Gaps 1;
 OY 1 YRLIRYRLAIR 12
 DB 1 YRLAIR-RIALR 11
 ||||| :|:|
 RESULT 4
 AAW33798
 ID AAW33798 standard; peptide; 12 AA.
 XX
 AC AAW33798;
 XX
 DT 19-JUN-1998 (first entry)
 XX
 DE Peptide B2702.84-79/79-84 tested for immunomodulating activity.
 XX
 KW Immunomodulating dimer; immunosuppressant drug; CTL activation;
 KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;
 KW rejection.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 PN WO9744351-A1.
 XX
 PD 27-NOV-1997.
 XX
 PF 22-MAY-1997; 97WO-US08689.
 XX
 PR 24-MAY-1996; 96US-0653294.
 XX
 XX (STRD) UNIV LELAND STANFORD JUNIOR.
 XX
 PI Beulow R, Clayberger C, Krensky AM;
 XX
 DR WPI; 1998-086530/08.
 XX
 PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B
 PT alpha-1 domain, used for preventing rejection of transplants or
 PT treating autoimmune diseases
 XX
 PS Example 1; Page 19; 41pp; English.
 XX
 CC Peptides AAW33784-98 and AAW33778-9 were assayed for their
 CC immunomodulating activity. A peptide-type compound or variant is claimed
 CC which has immunomodulating activity, including the N-terminal acylated
 CC and/or C-terminal amidated or esterified forms of up to 60 amino acids,
 CC where the peptide-type compound comprises the formula: A-B, where A, B =
 CC (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or V; aa77 =
 CC D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a
 CC small amino acid; aa82 = R or L; aa83 = G or R; and aa
 CC acid. The sequence in the brackets may optionally be absent or truncated
 CC at any peptide type bond within the brackets. The compounds comprise
 CC amino acid sequences related to a Class I HLA-B alpha domain (positions
 CC 79-84). They can be used to inhibit cytotoxic T-lymphocytes (CTL) from
 CC undesirably attacking cells in a host or in vitro. They can also be
 CC used in combination with antigenic peptides or proteins of interest to
 CC activate CTLs. They can also inhibit the proliferation of T cells in
 CC response to anti-CD3. The peptide can be used for preventing rejection
 CC of transplants or for treating autoimmune diseases, e.g. diabetes,
 CC rheumatoid arthritis and lupus erythematosus. The products can also be
 CC used for detection and diagnosis.

SQ Sequence 12 AA;
 Query Match 54.3%; Score 31.5; DB 19; Length 12;
 Best Local Similarity 66.7%; Pred. No. 14;
 Matches 8; Conservative 2; Mismatches 1; Indels 1; Gaps 1;
 OY 1 YRLIRYRLAIR 12
 DB 1 YRLAIR-RIALR 11
 ||||| :|:|
 RESULT 5
 AAW33799
 ID AAW33799 standard; peptide; 12 AA.
 XX
 AC AAW33799;
 XX
 DT 19-JUN-1998 (first entry)
 XX
 DE Immunomodulating dimer peptide #3.
 XX
 KW Immunomodulating dimer; immunosuppressant drug; CTL activation;
 KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;
 KW rejection.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 PN WO9744351-A1.
 XX
 PD 27-NOV-1997.
 XX
 PF 22-MAY-1997; 97WO-US08689.
 XX
 PR 24-MAY-1996; 96US-0653294.
 XX
 XX (STRD) UNIV LELAND STANFORD JUNIOR.
 XX
 PI Beulow R, Clayberger C, Krensky AM;
 XX
 DR WPI; 1998-086530/08.
 XX
 PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B
 PT alpha-1 domain, used for preventing rejection of transplants or
 PT treating autoimmune diseases
 XX
 PS Claim 17; Page 35; 41pp; English.
 XX
 CC This sequence represents a specifically claimed immunomodulating
 CC dimer peptide of the invention. A peptide-type compound or variant is
 CC claimed which has immunomodulating activity, including the N-terminal
 CC acylated and/or C-terminal amidated or esterified forms of up to 60
 CC amino acids, where the peptide-type compound comprises the formula: A-B,
 CC where A, B = (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or
 CC V; aa77 = D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a
 CC hydrophobic or small amino acid; aa82 = R or L; aa83 = G or R; and aa
 CC represents amino acid. The sequence in the brackets may optionally be
 CC absent or truncated at any peptide type bond within the brackets. The
 CC compounds comprise amino acid sequences related to a Class I HLA-B
 CC alpha domain (positions 79-84). They can be used to inhibit cytotoxic
 CC T-lymphocytes (CTL) from undesirably attacking cells in a host or in
 CC vitro. They can also be used in combination with antigenic peptides or
 CC proteins of interest to activate CTLs. They can also inhibit the
 CC proliferation of T cells in response to anti-CD3. The peptide can be
 CC used for preventing rejection of transplants or for treating autoimmune
 CC diseases, e.g. diabetes, rheumatoid arthritis and lupus erythematosus.
 CC The products can also be used for detection and diagnosis.
 XX
 SQ Sequence 12 AA;
 Query Match 54.3%; Score 31.5; DB 19; Length 12;
 Best Local Similarity 66.7%; Pred. No. 14;
 Matches 8; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

OY 1 YRLIRYRLAIR 12
||| | | | | |
Db 1 YRLAIR-R1ALR 11

RESULT 6
ABB29624
ID ABB29624 standard; Peptide; 28 AA.

AC ABB29624;

DT 01-FEB-2002 (first entry)

XX Peptide #2275 encoded by breast cell single exon nucleic acid probe.

DE Human; microarray; single exon probe; gene expression; breast;
KW disease; cancer.

XX Homo sapiens.

PN WO200157271-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00662.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SQ, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-496933/54.

XX New spatially-addressable set of single exon nucleic acid probes,
PT useful for measuring gene expression in sample derived from human
PS breast, comprises number of single exon nucleic acid probes -

XX Claim 27; SEQ ID NO 12592; 327pp + sequence listing; English.

XX The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human breast and BT 474 cells. The method involves contacting
CC the probes with a collection of detectably labelled nucleic acids
CC derived from mRNA of human breast, and then measuring the label
CC bound to each probe of the microarray. The probes are useful for
CC verifying the expression of regions of genomic DNA predicted to
CC encode proteins. They are useful for gene discovery, and for
CC determining predisposition and/or prognosing breast disease. Gene
CC expression analysis is useful for assessing the toxicity of chemical
CC agents on cells. The microarray of this invention presents a far greater
CC diversity of probes for measuring gene expression, with far less bias
CC than expressed sequence tag microarrays. The method is suitable for
CC rapid production of functional information from genomic sequence. The
CC present sequence is a peptide encoded by a single exon nucleic acid
CC probe of the invention.

CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 28 AA;

XX

Query Match 53.4%; Score 31; DB 22; Length 28;
Best Local Similarity 70.0%; Pred. No. 43;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 RLIRYRLAI 11
||| | | | | |
Db 1 RLIRYSLAV 10

RESULT 7
ABB34801
ID ABB34801 standard; Peptide; 28 AA.

AC ABB34801;

DT 04-FEB-2002 (first entry)

XX Peptide #2307 encoded by human foetal liver single exon probe.

XX Human; foetal liver; gene expression; single exon nucleic acid probe.

XX Homo sapiens.

PN WO200157277-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00669.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SQ, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-483447/52.

XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human fetal liver -

XX Claim 27; SEQ ID NO 27436; 639pp + sequence listing; English.

XX The invention relates to a single exon nucleic acid probe for
CC measuring human gene expression in a sample derived from human foetal
CC liver. The single exon nucleic acid probes may be used for predicting,
CC measuring and displaying gene expression in samples derived from human
CC fetal liver. The present sequence is a peptide encoded by a single exon
CC nucleic acid probe of the invention.

CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 28 AA;

XX

Query Match 53.4%; Score 31; DB 22; Length 28;
Best Local Similarity 70.0%; Pred. No. 43;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 RLIRYRLAI 11
||| | | | | |
Db 1 RLIRYSLAV 10

RESULT 8
ABB20217
ID ABB20217 standard; Protein; 28 AA.

AC ABB20217;

DT 23-JAN-2002 (first entry)

DE Protein #2216 encoded by probe for measuring heart cell gene expression.
 KW Human; gene expression; heart; microarray; vascular system;
 KW cardiovascular disease; hypertension; cardiac arrhythmia;
 KW congenital heart disease.
 XX Homo sapiens.
 OS
 XX
 XX WO200157274-A2.
 PN
 XX
 XX 09-AUG-2001.
 PD
 XX
 XX 30-JAN-2001; 2001WO-US00666.
 PF
 XX
 XX 04-FEB-2000; 2000US-0180312.
 PR
 XX 26-MAY-2000; 2000US-0207456.
 PR
 XX 30-JUN-2000; 2000US-0608408.
 PR
 XX 03-AUG-2000; 2000US-0632366.
 PR
 XX 21-SEP-2000; 2000US-0234687.
 PR
 XX 27-SEP-2000; 2000US-0236359.
 PR
 XX 04-OCT-2000; 2000GB-0024263.
 XX
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 PI
 XX WPI; 2001-488899/53.
 DR
 XX Single exon nucleic acid probes for analyzing gene expression in human
 PT hearts -
 PT
 XX Claim 15; SEQ ID No 21987; 530pp; English.
 PS
 XX The present invention relates to single exon nucleic acid probes for
 CC measuring human gene expression in a sample derived from human heart (see
 CC ABA21535-ABA41105). The present sequence is a protein encoded by one such
 CC probe. The probes may be used for predicting, measuring and displaying
 CC gene expression in samples derived from the human heart via microarrays.
 CC By measuring gene expression, the probes are useful for predicting,
 CC diagnosing, grading, staging, monitoring and prognosing diseases of the
 CC human heart and vascular system e.g. cardiovascular disease,
 CC hypertension, cardiac arrhythmias and congenital heart disease.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 XX Sequence 28 AA;
 SQ
 Query Match 53.4%; Score 31; DB 22; Length 28;
 Best Local Similarity 70.0%; Pred. No. 43;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 2 RLLIYRLAI 11
 Db |||||
 1 RLLIFYS LAV 10
 RESULT 9
 AAM55603
 ID AAM55603 standard; Protein; 28 AA.
 AC
 XX
 XX AAM55603;
 XX
 XX 05-NOV-2001 (first entry)
 DT
 XX
 XX Human brain expressed single exon probe encoded protein SEQ ID NO: 27708.
 DE
 XX Human; brain expressed exon; gene expression analysis; probe;
 KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
 KW epilepsy; cancer.
 XX
 XX Homo sapiens.
 OS
 XX

PN WO200157275-A2.
 XX
 PD 09-AUG-2001.
 XX
 XX 30-JAN-2001; 2001WO-US00667.
 PF
 XX
 XX 04-FEB-2000; 2000US-0180312.
 PR
 XX 26-MAY-2000; 2000US-0207456.
 PR
 XX 30-JUN-2000; 2000US-0608408.
 PR
 XX 03-AUG-2000; 2000US-0632366.
 PR
 XX 21-SEP-2000; 2000US-0234687.
 PR
 XX 27-SEP-2000; 2000US-0236359.
 PR
 XX 04-OCT-2000; 2000GB-0024263.
 XX
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 PI
 XX WPI; 2001-483446/52.
 DR
 XX Single exon nucleic acid probes for analyzing gene expression in human
 PT brains -
 PT
 XX Example 4; SEQ ID NO: 27708; 650pp + Sequence Listing; English.
 PS
 XX The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC brain. They can be used to measure gene expression in brain cell samples,
 CC which may enable the diagnosis and improved treatment of nervous system
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
 CC epilepsy and cancers. The present sequence is a protein encoded by one of
 CC the probes of the invention.
 XX
 XX Sequence 28 AA;
 SQ
 Query Match 53.4%; Score 31; DB 22; Length 28;
 Best Local Similarity 70.0%; Pred. No. 43;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 2 RLLIYRLAI 11
 Db |||||
 1 RLLIFYS LAV 10
 RESULT 10
 AAM67990
 ID AAM67990 standard; Protein; 28 AA.
 AC
 XX
 XX AAM67990;
 XX
 XX 06-NOV-2001 (first entry)
 DT
 XX
 XX Human bone marrow expressed probe encoded protein SEQ ID NO: 28296.
 DE
 XX Human; bone marrow expressed exon; gene expression analysis; probe;
 KW microarray; cancer; leukaemia; lymphoma; myeloma.
 KW
 XX
 XX Homo sapiens.
 OS
 XX
 XX WO200157276-A2.
 PN
 XX
 XX 09-AUG-2001.
 PD
 XX
 XX 30-JAN-2001; 2001WO-US00668.
 PF
 XX
 XX 04-FEB-2000; 2000US-0180312.
 PR
 XX 26-MAY-2000; 2000US-0207456.
 PR
 XX 30-JUN-2000; 2000US-0608408.
 PR
 XX 03-AUG-2000; 2000US-0632366.
 PR
 XX 21-SEP-2000; 2000US-0234687.
 PR
 XX 27-SEP-2000; 2000US-0236359.
 PR
 XX 04-OCT-2000; 2000GB-0024263.
 XX

PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488901/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human bone marrow -
XX
XX Example 4; SEQ ID NO: 28296; 658bp + sequence listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukemia and myeloma. The present sequence is a
CC protein encoded by one of the probes of the invention.
XX
SQ Sequence 28 AA;

Query Match 53.4%; Score 31; DB 22; Length 28;
Best Local Similarity 70.0%; Pred. No. 43;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 RLIRYRLAI 11
DB 1 RLIRYSLAV 10

RESULT 11
AAM15805
ID AAM15805 standard; Protein; 28 AA.
XX
XX AAM15805;
XX
XX 12-OCT-2001 (first entry)
XX
XX Peptide #2239 encoded by probe for measuring cervical gene expression.
XX
XX Probe; human; microarray; gene expression; cervical epithelial cell;
XX cervical cancer.
XX
XX Homo sapiens.
XX
XX WO200157278-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00670.
XX
XX 04-FEB-2000; 2000US-0180312.
XX
XX 26-MAY-2000; 2000US-0207456.
XX
XX 30-JUN-2000; 2000US-0608408.
XX
XX 03-AUG-2000; 2000US-0632366.
XX
XX 21-SEP-2000; 2000US-0234687.
XX
XX 27-SEP-2000; 2000US-0236359.
XX
XX 04-OCT-2000; 2000GB-0024263.
XX
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488901/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human cervical epithelial cells -
XX
XX Claim 27; SEQ ID No 20631; 487bp; English.
XX
XX The present invention relates to human single exon nucleic acid probes
CC (SENP: see A110068-A118459). The present sequence is a peptide encoded
CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs
CC can be used to produce a single exon microarray, which can be used for

CC measuring human gene expression in a sample derived from human cervical
CC epithelial cells. By measuring gene expression, the probes are therefore
CC useful in grading and/or staging of diseases of the cervix, notably
CC cervical cancer.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 28 AA;

Query Match 53.4%; Score 31; DB 22; Length 28;
Best Local Similarity 70.0%; Pred. No. 43;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 RLIRYRLAI 11
DB 1 RLIRYSLAV 10

RESULT 12
AAM28314
ID AAM28314 standard; Protein; 28 AA.
XX
XX AAM28314;
XX
XX 17-OCT-2001 (first entry)
XX
XX Peptide #2351 encoded by probe for measuring placental gene expression.
XX
XX Probe; microarray; human; placenta; antenatal diagnosis;
XX genetic disorder.
XX
XX Homo sapiens.
XX
XX WO200157272-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00663.
XX
XX 04-FEB-2000; 2000US-0180312.
XX
XX 26-MAY-2000; 2000US-0207456.
XX
XX 30-JUN-2000; 2000US-0608408.
XX
XX 03-AUG-2000; 2000US-0632366.
XX
XX 21-SEP-2000; 2000US-0234687.
XX
XX 27-SEP-2000; 2000US-0236359.
XX
XX 04-OCT-2000; 2000GB-0024263.
XX
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488901/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human placenta -
XX
XX Claim 27; SEQ ID No 28583; 654bp; English.
XX
XX The present invention relates to single exon nucleic acid probes (SENP:
CC see A113115-A1157546). The present sequence is a microarray for
CC such probe. The probes are useful for producing a microarray for
CC predicting, measuring and displaying gene expression in samples derived
CC from human placenta. The probes are useful for antenatal diagnosis of
CC human genetic disorders.
XX
SQ Sequence 28 AA;

Query Match 53.4%; Score 31; DB 22; Length 28;
Best Local Similarity 70.0%; Pred. No. 43;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 RLIRYRLAI 11

```

Db      1 RLLIFYSLAV 10
RESULT 13
AAM03543
ID AAM03543 standard; Protein; 28 AA.
XX
AC AAM03543;
XX
DT 09-OCT-2001 (first entry)
XX
DE Peptide #2225 encoded by probe for measuring breast gene expression.
XX
KW Probe; human; breast disease; breast cancer; development disorder;
KW inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX
OS Homo sapiens.
XX
FN WO200157270-A2.
XX
PD 09-AUG-2001.
XX
PF 29-JAN-2001; 2001WO-US00661.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234487.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
WPI; 2001-476286/51.
XX
DR Novel single exon nucleic acid probe used to measuring gene expression
PT in a human breast -
PT
PS Claim 27; SEQ ID No 12283; 322pp; English.
XX
CC The present invention relates to novel single exon nucleic acid probes
CC (see AAI00010-AAI10067). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for measuring human gene expression in
CC a human breast sample, where the probe hybridises at high stringency to a
CC nucleic acid expressed in the human breast. The probes are useful for
CC predicting, diagnosing, grading, staging, monitoring and prognosing
CC diseases of the human breast, particularly those diseases with polygenic
CC aetiology. The diseases include: breast cancer, disorders of development,
CC inflammatory diseases of the breast, fibrocystic changes, proliferative
CC breast disease and non-carcinoma tumours.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 28 AA;
Query Match 53.4%; Score 31; DB 22; Length 28;
Best Local Similarity 70.0%; Pred. No. 43;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 2 RLLIYRLAI 11
Db      1 RLLIFYSLAV 10
RESULT 14
ABG37525
ID ABG37525 standard; Peptide; 28 AA.
XX

```

```

AC ABG37525;
XX
DT 19-AUG-2002 (first entry)
XX
DE Human peptide encoded by genome-derived single exon probe SEQ ID 27190.
XX
KW Human; single exon probe; asthma; lung cancer; COPD; ILD;
KW chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesia; pulmonary hypertension;
KW hyaline membrane disease.
XX
OS Homo sapiens.
XX
FN WO200186003-A2.
XX
PD 15-NOV-2001.
XX
PF 30-JAN-2001; 2001WO-US00665.
XX
PR 04-FEB-2000; 2000US-180312P.
PR 26-MAY-2000; 2000US-207456P.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-234587P.
PR 27-SEP-2000; 2000US-236359P.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
WPI; 2002-114183/15.
XX
DR Spatially-addressable set of single exon nucleic acid probes, used to
PT measure gene expression in human lung samples -
PT
PS Claim 27; SEQ ID No 27190; 634pp; English.
XX
CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human lung comprising single exon nucleic acid probes having one of
CC 12614 nucleic acid sequences mentioned in the specification, or their
CC complements or the 12387 open reading frames derived from the 12614
CC probes. Also included are a microarray comprising the novel set of
CC probes; the novel set of probes which hybridise at high stringency to a
CC nucleic acid expressed in the human lung; measuring gene expression in a
CC sample derived from human lung, comprising (a) contacting the array with
CC a collection of detectably labeled nucleic acids derived from human lung
CC mRNA, and (b) measuring the label detectably bound to each probe of
CC the array; identifying exons in a eukaryotic genome, comprising
CC (a) algorithmically predicting at least one exon from genomic sequences
CC of the eukaryote; and (b) detecting specific hybridisation of detectably
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray; assigning exons to a single gene,
CC comprising (a) identifying exons from genomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridisation to a single exon
CC microarrays having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene
CC expression analysis, and for identifying exons in a gene, particularly
CC using human lung derived mRNA and for the study of lung diseases
CC such as asthma, lung cancer, chronic obstructive pulmonary disease
CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,

```

CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
 CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
 CC pulmonary alveolar proteinosis, Karsagen syndrome, fibrocystic
 CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
 CC and hyaline membrane disease. The present sequence is a peptide/protein
 CC encoded by a single exon probe of the invention.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

CC Sequence 28 AA;

Query Match 53.4%; Score 31; DB 23; Length 28;

Best Local Similarity 70.0%; Pred. No. 43;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 RLIRYRLAI 11
 Db 1 RLIRYSLAV 10

RESULT 15

AA001990
 ID AA001990 standard; Protein; 50 AA.

XX AA001990;

XX 06-NOV-2001 (first entry)

DE Human polypeptide SEQ ID NO 15882.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorders; arthritis; inflammation.

XX Homo sapiens.

OS WO200164835-A2.

XX 07-SEP-2001.

XX 26-FEB-2001; 2001WO-US04927.

XX 28-FEB-2000; 2000US-0515126.

XX 18-MAY-2000; 2000US-0577409.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-514838/56.

XX N-PSDB; AAI81921.

XX Isolated nucleic acids and polypeptides, useful for preventing
 PT diagnosing and treating e.g. leukaemia, inflammation and immune
 PT disorders -

XX Claim 20; SEQ ID NO 15882; 1399pp + Sequence listing; English.

XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
 CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.
 CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 50 AA;

XX Query Match

Best Local Similarity 71.4%; Pred. No. 80;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 YRLIRY 7

Db 4 FRLIVRY 10

Search completed: December 19, 2002; 16:19:09
 Job time : 29.5882 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 19, 2002, 16:14:32 ; Search time 21.1765 Seconds
(without alignments)
116.760 Million cell updates/sec

Title: US-08-653-294C-42
Perfect score: 58
Sequence: 1 YRLIRYLRLAIR 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 45785

Minimum DB seq length: 0
Maximum DB seq length: 60

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_21.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mmc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	33	56.9	35	12 Q65737	Q65737 bluetongue
2	30	51.7	43	16 Q8XG89	Q8XG89 salmonella
3	29	50.0	53	12 Q91FL0	Q91FL0 chilo iride
4	28	48.3	39	16 Q8X2E8	Q8X2E8 escherichia
5	28	48.3	39	16 P75731	P75731 escherichia
6	28	48.3	45	13 Q9PW17	Q9PW17 catla catla
7	28	48.3	58	16 Q8ZNG1	Q8ZNG1 salmonella
8	27	46.6	50	12 Q91FZ9	Q91FZ9 chilo iride
9	27	46.6	53	16 Q9KCES	Q9KCES bacillus ha
10	27	46.6	54	17 Q9YDU1	Q9YDU1 aeropyrum p
11	27	46.6	57	12 Q9QJX3	Q9QJX3 mumps virus
12	27	46.6	57	12 Q9U4L2	Q9U4L2 mumps virus
13	27	46.6	57	12 Q8V5G5	Q8V5G5 mumps virus
14	27	46.6	57	16 Q8YKB6	Q8YKB6 anabaena sp
15	26	44.8	31	1 Q8X254	Q8X254 halobacteri
16	26	44.8	37	2 Q9X9E2	Q9X9E2 uncultured

17	26	44.8	37	2 Q9X9E3	Q9X9E3 uncultured
18	26	44.8	37	2 Q9X9E4	Q9X9E4 uncultured
19	26	44.8	37	2 Q9X9E5	Q9X9E5 uncultured
20	26	44.8	37	2 Q9X9E6	Q9X9E6 uncultured
21	26	44.8	37	2 Q9X9E7	Q9X9E7 uncultured
22	26	44.8	37	2 Q9X9E8	Q9X9E8 uncultured
23	26	44.8	37	2 Q9X9E9	Q9X9E9 uncultured
24	26	44.8	37	2 Q9X9F0	Q9X9F0 uncultured
25	26	44.8	45	15 Q99BL0	Q99BL0 human immun
26	26	44.8	45	15 Q8UTB6	Q8UTB6 human immun
27	26	44.8	49	2 Q8VVT4	Q8VVT4 staphylococ
28	26	44.8	50	16 Q50691	Q50691 borrelia bu
29	26	44.8	54	6 Q95M82	Q95M82 equus cabal
30	26	44.8	55	8 Q78795	Q78795 pytaiaella l
31	26	44.8	57	4 Q9P1C8	Q9P1C8 homo sapien
32	26	44.8	57	12 P89006	P89006 mumps virus
33	26	44.8	57	12 P89008	P89008 mumps virus
34	26	44.8	57	12 P89009	P89009 mumps virus
35	26	44.8	57	12 P89013	P89013 mumps virus
36	26	44.8	57	12 Q98349	Q98349 mumps virus
37	26	44.8	57	12 Q98350	Q98350 mumps virus
38	26	44.8	57	12 Q9WAH4	Q9WAH4 mumps virus
39	26	44.8	57	12 Q9QJX0	Q9QJX0 mumps virus
40	26	44.8	57	12 Q9QJW9	Q9QJW9 mumps virus
41	26	44.8	57	12 Q9QJW8	Q9QJW8 mumps virus
42	26	44.8	57	12 Q9QEV1	Q9QEV1 mumps virus
43	26	44.8	57	12 Q9QEV0	Q9QEV0 mumps virus
44	26	44.8	57	12 Q9QEU8	Q9QEU8 mumps virus
45	26	44.8	57	12 Q9QEU7	Q9QEU7 mumps virus

ALIGNMENTS

RESULT 1

Q65737 PRELIMINARY; PRT; 35 AA.
ID AC Q65737;
DC 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-NOV-1998 (Tremblrel. 08, Last annotation update)
DE Serotype-specific antigen (S2) RNA, 5' end (Fragment).
OS Bluetongue virus.
OC Viruses; dsRNA viruses; Reoviridae; Orbivirus.
OX NCBI_TaxID=12591;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88072081; PubMed=2825414;
RA Huismans H., Cloete M., le Roux A.;
RT "The genetic relatedness of a number of individual cognate genes of
RT viruses in the bluetongue and closely related serogroups.";
RL Virology 161:421-428(1987).
DR EMBL; M22442; AAA42841.1; -.
FT NON TER 35
SQ SEQUENCE 35 AA; 4186 MW; C7E02BAA0510DFF5 CRC64;

Query Match 56.9%; Score 33; DB 12; Length 35;

Best Local Similarity 66.7%; Pred. No. 28;

Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 YRLIRYLRLAIR 12

Db 16 YSLNHYPLAIR 27

RESULT 2

Q8XG89 PRELIMINARY; PRT; 43 AA.
ID Q8XG89;
DC 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Putative inner membrane protein (Hypothetical protein STV3241).

GN YOG8 OR STM3087 OR STY3241.
OS Salmomella typhimurium, and
OS Salmomella typh.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmomella.
OX NCBI_TaxID=602, 601;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhimurium; STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porrolik S., Ali J., Dante M., Du P., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Sconking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmomella enterica serovar Typhimurium
LT2.";
RL Nature 413:852-856(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typh; STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Main J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor J.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Feltham T., Hamlin N., Haque A., Hien T.T., Holtroyd S., Jorgle K.,
RA Krogsh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrett B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmomella
enterica serovar Typh CT18.";
RL Nature 413:848-852(2001).
DR EMBL; AE008842; AAL21962.1; -;
DR EMBL; AL627277; CAD02912.1; -;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 43 AA; 4954 MW; 5489F054F09E2277 CRC64;

Query Match 51.7%; Score 30; DB 16; Length 43;
Best Local Similarity 63.6%; Pred. No. 1.3e+02;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 YRLIRYRLAI 11
Db 22 YGLSLRLRLAI 32

RESULT 3
ID Q91FLO PRELIMINARY; PRT; 53 AA.
AC Q91FLO;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE 314L.
OS Chilo iridescent virus (CIV) (insect iridescent virus type 6).
OC Viruses; dsDNA viruses, no RNA stage; Iridoviridae; Iridovirus.
OX NCBI_TaxID=10488;
RN [1]
RP SEQUENCE FROM N.A.
RA Delius H., Darai G., Fluegel R.M.;
RT "DNA analysis of insect iridescent virus 6: evidence for circular
permutation and terminal redundancy.";
RL J. Virol. 49:609-614(1984).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=86174607; PubMed=3959991;
RA Lorbacher de Ruiz V., Gelderblom H., Hofmann W., Darai G.;
RT "Insect iridescent virus type 6 induced toxic degenerative hepatitis
in mice.";
RL Med. Microbiol. Immunol. 175:43-53(1986).
RN [3]
RP SEQUENCE FROM N.A.

RX MEDLINE=87321126; PubMed=2820141;
RA Schutzieler P., Soltau J.B., Fischer M., Reissner H., Scholz J.,
RA Delius H., Darai G.;
RT "Molecular cloning and physical mapping of the genome of insect
iridescent virus type 6: further evidence for circular permutation of
the viral genome.";
RL Virology 160:66-74(1987).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=89073752; PubMed=3201750;
RA Fischer M., Schutzieler P., Delius H., Darai G.;
RT "Identification and characterization of the repetitive DNA element in
the genome of insect iridescent virus type 6.";
RL Virology 167:485-496(1988).
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=92196966; PubMed=1549908;
RA Handermann M., Schutzieler P., Rosen-Wolff A., Raab K., Sonntag K.C.,
RA Darai G.;
RT "Characterization of the third origin of DNA replication of the genome
of insect iridescent virus type 6.";
RL Virus Genes 6:333-342(1992).
RN [6]
RP SEQUENCE FROM N.A.
RX MEDLINE=93118242; PubMed=1475907;
RA Sonntag K.C., Darai G.;
RT "Identification of the third origin of DNA replication of the genome
of insect iridescent virus type 6.";
RL Virus Genes 6:333-342(1992).
RN [7]
RP SEQUENCE FROM N.A.
RX MEDLINE=93260401; PubMed=8492091;
RA Stomwasser R., Raab K., Schutzieler P., Janssen W., Darai G.;
RT "Identification of the gene encoding the major capsid protein of
insect iridescent virus type 6 by polymerase chain reaction.";
RL J. Gen. Virol. 74:873-879(1993).
RN [8]
RP SEQUENCE FROM N.A.
RX MEDLINE=94167241; PubMed=8121799;
RA Schutzieler P., Hug M., Handermann M., Janssen W., Koonin E.V.,
RA Delius H., Darai G.;
RT "Identification of genes encoding zinc finger proteins, non-histone
chromosomal HMG protein homologue, and a putative GTP phosphohydrolase
in the genome of Chilo iridescent virus.";
RL Nucleic Acids Res. 22:158-166(1994).
RN [9]
RP SEQUENCE FROM N.A.
RX MEDLINE=94353641; PubMed=8073636;
RA Sonntag K.C., Schutzieler P., Koonin E.V., Darai G.;
RT "Chilo iridescent virus encodes a putative helicase belonging to a
distinct family within the 'DEAD/H' superfamily: implications for the
evolution of large DNA viruses.";
RL Virus Genes 8:151-158(1994).
RN [10]
RP SEQUENCE FROM N.A.
RX MEDLINE=95213160; PubMed=7698884;
RA Sonntag K.C., Schutzieler P., Janssen W., Darai G.;
RT "Identification of the primary structure and the coding capacity of
the genome of insect iridescent virus type 6 between the genome
coordinates 0.310 and 0.347 (7990 bp).";
RL Intervirology 37:287-297(1994).
RN [11]
RP SEQUENCE FROM N.A.
RX MEDLINE=94292906; PubMed=8021587;
RA Schutzieler P., Sonntag K.C., Muller M., Janssen W., Bugart J.J.,
RA Koonin E.V., Darai G.;
RT "Insect iridescent virus type 6 encodes a polypeptide related to the
largest subunit of eukaryotic RNA polymerase II.";
RL J. Gen. Virol. 75:1557-1567(1994).
RN [12]
RP SEQUENCE FROM N.A.
RX MEDLINE=98141693; PubMed=9482589;
RA Bahr U., Tidona C.A., Darai G.;

RT "The DNA sequence of Chilo iridescent virus between the genome
RT coordinates 0.101 and 0.391; similarities in coding strategy between
RT insect and vertebrate iridoviruses.";
RL Virus Genes 15:235-245(1997).
RN [13]

RP SEQUENCE FROM N.A.
RX MEDLINE=99125223; PubMed=9926400;
RA Muller K., Tidona C.A., Bahr U., Darai G.;
RT "Identification of a thymidylate synthase gene within the genome of
RT Chilo iridescent virus.";
RL Virus Genes 17:243-258(1998).
RN [14]

RP SEQUENCE FROM N.A.
RX MEDLINE=99383793; PubMed=10456793;
RA Muller K., Tidona C.A., Darai G.;
RT "Identification of a gene cluster within the genome of Chilo
RT iridescent virus encoding enzymes involved in viral DNA replication
RT and processing.";
RL Virus Genes 18:243-264(1999).
RN [15]

RP SEQUENCE FROM N.A.
RX MEDLINE=21342589; PubMed=11448171;
RA Jakob N.J., Muller K., Bahr U., Darai G.;
RT "Analysis of the First Complete DNA Sequence of an Invertebrate
RT Iridovirus: Coding Strategy of the Genome of Chilo Iridescent Virus.";
RL Virology 286:182-196(2001).
RN [16]

RP SEQUENCE FROM N.A.
RA Jakob N.J., Mueller K., Bahr U., Darai G.;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF303741; AK82175.1; -
SQ SEQUENCE 53 AA; 6610 MW; 6C206BDEFA37C3B8 CRC64;

Query Match 50.0%; Score 29; DB 12; Length 53;
Best Local Similarity 45.5%; Pred. No. 2.4e+02;
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 YRLIYRLAI 11
: : : : :
DB 27 FKLNLKYLKI 37

RESULT 4
Q8X2E8 PRELIMINARY; PRT; 39 AA.
AC Q8X2E8;
DT 01-MAR-2002 (TREMELrel. 20, Created)
DT 01-MAR-2002 (TREMELrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMELrel. 20, Last annotation update)
DE Putative RNA.
GN ECS0703.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Okishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:111-22(2001).
DR EMBL; AF002552; BAB34126.1; -
SQ SEQUENCE 39 AA; 4864 MW; 8CA3F814F57CD6F1 CRC64;

Query Match 48.3%; Score 28; DB 16; Length 39;
Best Local Similarity 54.5%; Pred. No. 2.8e+02;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 YRLIYRLAI 11
: : : : :
DB 21 YQKPVYRLAI 31

RESULT 5
P75731 PRELIMINARY; PRT; 39 AA.
AC P75731;
DT 01-FEB-1997 (TREMELrel. 02, Created)
DT 01-FEB-1997 (TREMELrel. 02, Last sequence update)
DT 01-MAR-2002 (TREMELrel. 20, Last annotation update)
DE From BASES 685948 to 696574 (Section 60 of 400) of the complete
DE genome.
GN B0669.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
DR EMBL; AB000170; AAC73766.1; -
KW Complete proteome.
SQ SEQUENCE 39 AA; 4865 MW; 8CAB5A9ED57CD6F1 CRC64;

Query Match 48.3%; Score 28; DB 16; Length 39;
Best Local Similarity 54.5%; Pred. No. 2.8e+02;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 YRLIYRLAI 11
: : : : :
DB 21 YQKPVYRLAI 31

RESULT 6
Q9PW17 PRELIMINARY; PRT; 45 AA.
AC Q9PW17;
DT 01-MAY-2000 (TREMELrel. 13, Created)
DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMELrel. 20, Last annotation update)
DE Hypothetical 5.3 kDa protein (fragment).
OS Catla catla (catla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Catla.
OX NCBI_TaxID=72446;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SPLEEN;
RA Ohri S., Vashishtha A., Dixit A.;
RT "Catla catla highly AT rich expressed sequences.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ249234; CAB53862.1; -
KW Hypothetical protein.
FT NON_TER 1
FT NON_TER 45
SQ SEQUENCE 45 AA; 5294 MW; 015ECB8ADB55BC0 CRC64;

Query Match 48.3%; Score 28; DB 13; Length 45;
Best Local Similarity 44.4%; Pred. No. 3.2e+02;
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 4 LIRYLAI 12
: : : : :
DB 4 LIRYLAI 12

```

Db      16  IYFRMAYR 24

RESULT 7
08ZNG1 PRELIMINARY; PRT; 58 AA.
ID 08ZNG1;
AC 08ZNG1;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DE Putative cytoplasmic protein.
GN STM2288.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spaeth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2."
RL Nature 413:852-856(2001).
DR EMBL; AE008802; AL21189.1;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 58 AA; 6709 MW; FAEAC712029A58C CRC64;

Query Match 48.3%; Score 28; DB 16; Length 58;
Best Local Similarity 85.7%; Pred. No. 4.1e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      6  RYRLAIR 12
Db      16  RYSLAIR 22

RESULT 8
091FZ9 PRELIMINARY; PRT; 50 AA.
ID 091FZ9;
AC 091FZ9;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE 152R.
OS Chilo iridescent virus (CIV) (insect iridescent virus type 6).
OC Viruses; dsDNA viruses, no RNA stage; Iridoviridae; Iridovirus.
OX NCBI_TaxID=10488;
RN [1]
RP SEQUENCE FROM N.A.
RA Delius H., Darai G., Fluegel R.M.;
RT "DNA analysis of insect iridescent virus 6: evidence for circular
RT permutation and terminal redundancy."
RL J. Virol. 49:609-614(1984).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=86174607; PubMed=3959991;
RA Lotacher de Ruiz H., Gelderblom H., Hofmann W., Darai G.;
RT "Insect iridescent virus type 6 induced toxic degenerative hepatitis
RT in mice."
RL Med. Microbiol. Immunol. 175:43-53(1986).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=87321126; PubMed=2820141;
RA Schittzler P., Soltan U.B., Fischer M., Reiser H., Scholz J.,
RA Delius H., Darai G.;
RT "Molecular cloning and physical mapping of the genome of insect
RT iridescent virus type 6: further evidence for circular permutation of
RT the viral genome."

Virology 160:66-74(1987).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=89073752; PubMed=3201750;
RA Fischer M., Schittzler P., Delius H., Darai G.;
RT "Identification and characterization of the repetitive DNA element in
RT the genome of insect iridescent virus type 6."
RL Virology 167:485-496(1988).
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=92196996; PubMed=1549908;
RA Handermann M., Schittzler P., Rosen-Wolff A., Raab K., Sonntag K.C.,
RA Darai G.;
RT "Identification and mapping of origins of DNA replication within the
RT DNA sequences of the genome of insect iridescent virus type 6."
RL Virus Genes 6:19-32(1992).
RN [6]
RP SEQUENCE FROM N.A.
RX MEDLINE=93118242; PubMed=1475907;
RA Sonntag K.C., Darai G.;
RT "Characterization of the third origin of DNA replication of the genome
RT of insect iridescent virus type 6."
RL Virus Genes 6:333-342(1992).
RN [7]
RP SEQUENCE FROM N.A.
RX MEDLINE=93260401; PubMed=8492091;
RA Stohwasser R., Raab K., Schittzler P., Janssen W., Darai G.;
RT "Identification of the gene encoding the major capsid protein of
RT insect iridescent virus type 6 by polymerase chain reaction."
RL J. Gen. Virol. 74:873-879(1993).
RN [8]
RP SEQUENCE FROM N.A.
RX MEDLINE=94167241; PubMed=8121799;
RA Schittzler P., Hng M., Handermann M., Janssen W., Koonin E.V.,
RA Delius H., Darai G.;
RT "Identification of genes encoding zinc finger proteins, non-histone
RT chromosomal HMG protein homologue, and a putative GTP phosphohydrolase
RT in the genome of Chilo iridescent virus."
RL Nucleic Acids Res. 22:158-166(1994).
RN [9]
RP SEQUENCE FROM N.A.
RX MEDLINE=94353641; PubMed=8073636;
RA Sonntag K.C., Schittzler P., Koonin E.V., Darai G.;
RT "Chilo iridescent virus encodes a putative helicase belonging to a
RT distinct family within the 'DEAD/H' superfamily: implications for the
RT evolution of large DNA viruses."
RL Virus Genes 8:151-158(1994).
RN [10]
RP SEQUENCE FROM N.A.
RX MEDLINE=95213160; PubMed=7698884;
RA Sonntag K.C., Schittzler P., Janssen W., Darai G.;
RT "Identification of the primary structure and the coding capacity of
RT the genome of insect iridescent virus type 6 between the genome
RT coordinates 0.310 and 0.347 (7990 bp)."
RL Intervirology 37:287-297(1994).
RN [11]
RP SEQUENCE FROM N.A.
RX MEDLINE=94292906; PubMed=8021587;
RA Schittzler P., Sonntag K.C., Muller M., Janssen W., Bugert J.J.,
RA Koonin E.V., Darai G.;
RT "Insect iridescent virus type 6 encodes a polypeptide related to the
RT largest subunit of eukaryotic RNA polymerase II."
RL J. Gen. Virol. 75:1557-1567(1994).
RN [12]
RP SEQUENCE FROM N.A.
RX MEDLINE=98141693; PubMed=9482589;
RA Bahr U., Tidona C.A., Darai G.;
RT "The DNA sequence of Chilo iridescent virus between the genome
RT coordinates 0.101 and 0.391; similarities in coding strategy between
RT insect and vertebrate Iridoviruses."
RL Virus Genes 15:235-245(1997).
RN [13]
RP SEQUENCE FROM N.A.

```

```

RX MEDLINE=99125223; PubMed=9926400;
RA Muller K., Tidona C.A., Bahr U., Darai G.;
RT "Identification of a thymidylate synthase gene within the genome of
RT Chilo iridescent virus.";
RL Virus Genes 17:243-258(1998).
RN [14]
RP SEQUENCE FROM N.A.
RX MEDLINE=99383793; PubMed=10456793;
RA Muller K., Tidona C.A., Darai G.;
RT "Identification of a gene cluster within the genome of Chilo
RT iridescent virus encoding enzymes involved in viral DNA replication
RT and processing.";
RL Virus Genes 18:243-264(1999).
RN [15]
RP SEQUENCE FROM N.A.
RX MEDLINE=21342589; PubMed=11448171;
RA Jakob N.J., Muller K., Bahr U., Darai G.;
RT "Analysis of the First Complete DNA Sequence of an Invertebrate
RT Iridovirus: Coding Strategy of the Genome of Chilo Iridescent Virus.";
RL Virology 286:182-196(2001).
RN [16]
RP SEQUENCE FROM N.A.
RA Jakob N.J., Mueller K., Bahr U., Darai G.;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF303741; AAK82033.1;
SQ SEQUENCE 50 AA; 5601 MW; C3390D2D93B68010 CRC64;

Query Match 46.6%; Score 27; DB 12; Length 50;
Best Local Similarity 40.0%; Pred. No. 5.5e+02;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 2 RLIRYLAI 11
DB 27 KILLIRLAV 36

RESULT 9
Q9KCE5 PRELIMINARY; PRT; 53 AA.
AC Q9KCE5;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein BH1626.
GN BH1626.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Bacillaceae; Bacillus.
OX NCBI_TaxID=86685;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331(2000).
DR EMBL: AP001512; BAB05345.1;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 53 AA; 6234 MW; 343F2BA731F5E63C CRC64;

Query Match 46.6%; Score 27; DB 16; Length 53;
Best Local Similarity 71.4%; Pred. No. 5.8e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 6 RYRLAIR 12
DB 5 RYRWIR 11

RESULT 10

```

```

Q9YDU1 PRELIMINARY; PRT; 54 AA.
AC Q9YDU1;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein APES036.
GN APES036.
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
OC Desulfurococaceae; Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K1;
RX MEDLINE=99310339; PubMed=10382966;
RA Kawarabayasi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
RT crenarchaeon, Aeropyrum pernix K1.";
RL DNA Res. 6:83-101(1999).
DR EMBL: AP000060; BAA79806.1;
DR InterPro; IPR001211; PhospholipaseA2.
DR PROSITE; PS00118; PA2_HIS; UNKNOWN 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 54 AA; 6140 MW; 7579866613EACF1C CRC64;

Query Match 46.6%; Score 27; DB 17; Length 54;
Best Local Similarity 71.4%; Pred. No. 5.9e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 LLIRYEL 9
DB 18 LLIRFEL 24

RESULT 11
Q9QJX3 PRELIMINARY; PRT; 57 AA.
AC Q9QJX3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Small hydrophobic protein.
OS Mumps virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Rubulaviruses.
OX NCBI_TaxID=11161;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CIRI/UK96;
RX MEDLINE=99367591; PubMed=10438373;
RA Jin L., Beard S., Brown D.W.G.;
RT "Genetic heterogeneity of mumps virus in the United Kingdom:
RT identification of two new genotypes.";
RL J. Infect. Dis. 180:829-833(1999).
DR EMBL: AF142761; AAD54491.1;
DR InterPro; IPR001477; SH.
DR Pfam; PF01445; SH; 1.
DR ProDom; PD001504; SH; 1.
SQ SEQUENCE 57 AA; 6874 MW; 89425CA2376B2FE2 CRC64;

Query Match 46.6%; Score 27; DB 12; Length 57;
Best Local Similarity 50.0%; Pred. No. 6.2e+02;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 LLIRYELAIR 12
DB 30 LTINYKTAVR 39

```

```

RESULT 12
O9J4L2
ID O9J4L2 PRELIMINARY; PRT; 57 AA.
AC O9J4L2;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Small hydrophobic protein.
GN SH.
OS Mumps virus (strain Jeryl-Lynn).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Rubulaviruses.
OX NCBI_TaxID=11168;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JERYL LYNN;
RX MEDLINE=20240069; PubMed=10775622;
RA Clarke D.K., Sidhu M.S., Johnson J.E., Udem S.A.;
RT "Rescue of mumps virus from cDNA.";
RL J. Virol. 74:4831-4838(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=JERYL LYNN;
RA Clarke D.K., Sidhu M.S., Johnson J.E., Udem S.A.;
RN Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RP [3]
RC SEQUENCE FROM N.A.
RA STRAIN=JERYL LYNN;
RA Ameis G., Chumakov K.;
RT "Quantitative analysis of sequence diversity in Jeryl Lynn strain of
RT mumps vaccine live.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF201473; AAF70394.1; -
DR EMBL: AF338106; AAK83229.1; -
DR InterPro: IPR001477; SH.
DR Pfam: PF01445; SH; 1.
DR ProDom: PD001504; SH; 1.
SQ SEQUENCE 57 AA; 6782 MW; 85468D646748213F CRC64;

Query Match 46.6%; Score 27; DB 12; Length 57;
Best Local Similarity 50.0%; Pred. No. 6.2e+02;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 3 LLIRYRLAIR 12
DB 30 LTINYKTAVR 39

RESULT 13
O8V5G5
ID O8V5G5 PRELIMINARY; PRT; 57 AA.
AC O8V5G5;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Small hydrophobic protein.
GN SH.
OS Mumps virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Rubulaviruses.
OX NCBI_TaxID=11161;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=V79-49863;
RA Tecle T., Bottiger B., Orvell C., Johansson B.;
RT "Characterization of two decades of temporal co-circulation of four
RT mumps virus genotypes in Denmark, including one new genotype.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF365883; AAL68449.1; -
DR InterPro: IPR001477; SH.
DR Pfam: PF01445; SH; 1.

```

```

DR ProDom: PD001504; SH; 1.
SQ SEQUENCE 57 AA; 6844 MW; C78821095BA08429 CRC64;

Query Match 46.6%; Score 27; DB 12; Length 57;
Best Local Similarity 50.0%; Pred. No. 6.2e+02;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 3 LLIRYRLAIR 12
DB 30 LTINYKTAVR 39

RESULT 14
O8YKB6
ID O8YKB6 PRELIMINARY; PRT; 57 AA.
AC O8YKB6;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Transposase.
GN ASR7385.
OS Anabaena sp. (strain PCC 7120).
OG Plasmid pCC7120alpha.
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneo T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriyuchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yanada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120.";
RL DNA Res. 8:205-213(2001).
DR EMBL: AP003601; BAB77143.1; -
KW Plasmid; Complete proteome.
SQ SEQUENCE 57 AA; 6537 MW; AEB77D1D4D747C40 CRC64;

Query Match 46.6%; Score 27; DB 16; Length 57;
Best Local Similarity 60.0%; Pred. No. 6.2e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 2 RLIRYRLAI 11
DB 38 RLIRRRRFGI 47

RESULT 15
O8X254
ID O8X254 PRELIMINARY; PRT; 31 AA.
AC O8X254;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Hypothetical 3.6 kDa protein (fragment).
OS Archaea; Euryarchaeota; Halobacteriales; Halobacteriaceae; Halobacterax.
OC Halobacteriaceae; Halobacterax.
OX NCBI_TaxID=2246;
RN [1]
RP SEQUENCE FROM N.A.
RA Bidle K.A.;
RT "Differential expression of genes influenced by changing salinity
RT using RNA arbitrarily primed PCR in the archaeal halophile, Halotferax
RT volcanii.";
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF454092; AAL57845.1; -
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 31 AA; 3609 MW; 682119419273C9D0 CRC64;

```

Query Match 44.8%; Score 26; DB 1; Length 31;
Best Local Similarity 45.5%; Pred. No. 5.5e+02;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 YRLIIRYRLAI 11
||| : :
Db 1 YRLIKHALEL 11

Search completed: December 19, 2002, 16:21:56
Job time : 22.1765 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 19, 2002, 16:03:12 ; Search time 5.47059 Seconds
(without alignments)
90.980 Million cell updates/sec

Title: US-08-653-294C-42
Perfect score: 58
Sequence: 1 YLLIRYRLAIR 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 5116

Minimum DB seq length: 0
Maximum DB seq length: 60

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	28.5	49.1	44	1 YIMB_BPPH1	P10431 bacteriophage
2	26	44.8	26	1 YFHA_KLEPN	P21710 klebsiella
3	26	44.8	36	1 PMV_PETMA	P80024 petromyzon
4	25	43.1	38	1 CRRP_CANPG	P81033 cancer pagu
5	25	43.1	38	1 VPU_HV1A2	P05949 human immun
6	25	43.1	40	1 Y382_TREPA	O83397 treponema p
7	25	43.1	46	1 Y718_ARCFU	O29540 archaeoglob
8	24.5	42.2	57	1 VSH_MUMPB	P28082 mumps virus
9	24	41.4	50	1 HSP1_PONPY	P35310 pongo pygma
10	24	41.4	54	1 LT01_HORVU	O42509 hordeum vul
11	24	41.4	54	1 LT02_HORVU	O9ard5 hordeum vul
12	24	41.4	56	1 LCRT_YERPS	O00932 versinia ps
13	23.5	40.5	49	1 Y195_BPT7	P03804 bacterioph
14	23	39.7	30	1 Y573_TREPA	O83583 treponema p
15	23	39.7	36	1 PYV_AMICA	P29205 amia calva
16	23	39.7	43	1 VPU_HV1C4	P08803 human immun
17	23	39.7	45	1 RL34_SYNY3	Q55004 synechocyst
18	23	39.7	45	1 Y00D_BPT4	O01434 bacterioph
19	23	39.7	46	1 YLCG_BP82	O37874 bacterioph
20	23	39.7	48	1 YQGB_ECOLI	P46877 escherichia
21	23	39.7	55	1 YCK1_CUSRE	P32035 cuscutea ref
22	23	39.7	57	1 VSH_MUMPA	P22111 mumps virus
23	23	39.7	57	1 VSH_MUMPL	P28083 mumps virus
24	23	39.7	57	1 VSH_MUMPM	P22112 mumps virus
25	23	39.7	57	1 VSH_MUMPR	P28087 mumps virus
26	22	37.9	27	1 YCK7_CYAPA	P48328 cyanophora
27	22	37.9	36	1 NEUY_PLG	P01304 sus scrofa
28	22	37.9	36	1 NEUY_SHEEP	P14765 ovis aries
29	22	37.9	38	1 PACA_URAJA	P81039 uranoscopus
30	22	37.9	46	1 YLCG_ECOLI	O47272 escherichia
31	22	37.9	49	1 HSP1_SAGIM	P24714 saginus im
32	22	37.9	50	1 HSP1_HYLLA	P35306 hylobates l
33	22	37.9	51	1 HSP1_ALOSE	P35302 alouatta se

RESULT 1

YIMB_BPPH1 STANDARD; PRT; 44 AA.
AC P10431;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 01-MAR-1989 (Rel. 10, Last annotation update)
DE Hypothetical immunity region protein 8.
OS Bacteriophage phi-105.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
OC Lambda-like viruses.
OC NCBI_TaxID=10717;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86056972; PubMed=3934047;
RT Cully D.F., Garro A.J.;
RT "Nucleotide sequence of the immunity region of Bacillus subtilis
bacteriophage phi 105: identification of the repressor gene and its
mRNA and protein products.";
RL Gene 38:153-164(1985).

ALIGNMENTS

34 22 37.9 55 1 ATP8_COTJA
35 22 37.9 57 1 VSH_MUMPI
36 22 37.9 57 1 VSH_MUMPE
37 22 37.9 57 1 VSH_MUMPK
38 21.5 37.1 57 1 VSH_MUMP2
39 21.5 37.1 57 1 VSH_MUMPA
40 21 36.2 9 1 BUK_CLOPA
41 21 36.2 22 1 RL18_HALME
42 21 36.2 23 1 RL18_HALVO
43 21 36.2 26 1 RL18_HALHA
44 21 36.2 30 1 RL18_HALCU
45 21 36.2 36 1 PAHO_ANSAN
P0682 coturnix co
P19719 mumps virus
P22109 mumps virus
P28086 mumps virus
P28084 mumps virus
P28085 mumps virus
P81337 clostridium
P50561 halobacteri
P50563 halobacteri
P05970 halobacteri
P06304 anser anser

Query Match 49.1%; Score 28.5; DB 1; Length 44;
Best Local Similarity 63.6%; Pred. No. 25;
Matches 7; Conservative 3; Mismatches 0; Indels 1; Gaps 1;

QY 3 LLIRY-RLAIR 12

Db 9 LLIRYRLALK 19

RESULT 2

YFHA_KLEPN STANDARD; PRT; 26 AA.
ID YFHA_KLEPN
AC P21710;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein in GLNB 5' region (Fragment).
GN YFHA.
OS Klebsiella pneumoniae.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Klebsiella.
OC NCBI_TaxID=573;

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M5a1;
RX MEDLINE=89201233; PubMed=2907369;
RA Holzel A., Merrick M.;
RT "Identification of the Klebsiella pneumoniae glbB gene: nucleotide
  sequence of wild-type and mutant alleles.";
RL Mol. Gen. Genet. 215:134-138(1988).
CC -1- FUNCTION: PROBABLE MEMBER OF A TWO-COMPONENT REGULATORY SYSTEM
  YFHA/YFHK.
CC -1- PTM: PHOSPHORYLATED BY YFHK (POTENTIAL).
CC -1- SIMILARITY: THE N-TERMINAL REGION IS SIMILAR TO THAT OF OTHER
  REGULATORY COMPONENTS OF SENSORY TRANSDUCTION SYSTEMS.
CC -1- SIMILARITY: THE CENTRAL REGION CONTAINS A SIGMA-54 FACTOR
  INTERACTION ATP-BINDING DOMAIN.
CC -1- SIMILARITY: TO EQUIVALENT PROTEIN IN E. COLI.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation-
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
  -----
CC EMBL; X14012; CAA32176.1; -.
DR PIR; S04376; S04376.
DR InterPro; IPR002078; Sig54_interact.
DR PROSITE; PS00675; SIGMA54_INTERACT_1; PARTIAL.
DR PROSITE; PS00676; SIGMA54_INTERACT_2; PARTIAL.
DR PROSITE; PS00688; SIGMA54_INTERACT_3; PARTIAL.
DR PROSITE; PS50045; SIGMA54_INTERACT_4; PARTIAL.
KW Hypothetical protein; Sensory transduction; Phosphorylation;
  Transcription regulation; DNA-binding; ATP-binding.
FT NON TER 1
SQ SEQUENCE 26 AA; 3198 MW; 510C8BBA43BF005 CRC64;

Query Match 44.8%; Score 26; DB 1; Length 26;
Best Local Similarity 55.6%; Pred. No. 45;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 YRLIRYRL 9
DB 11 YRLIRSHL 19

RESULT 3
PMV_PETMA STANDARD; PRT; 36 AA.
AC P80024;
DT 01-AUG-1991 (Rel. 19; Created)
DT 01-AUG-1991 (Rel. 19; Last sequence update)
DT 16-OCT-2001 (Rel. 40; Last annotation update)
DE Pancreatic polypeptide MY (PMV).
OS Petromyzon marinus (Sea lamprey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
  Petromyzontiformes; Petromyzontidae; Petromyzon.
OX NCBI_TaxID=7757;
RN [1]
RP SEQUENCE.
RC TISSUE=Intestine; PubMed=2070789;
RX MEDLINE=91301137; PubMed=2070789;
RA Conlon J.M., Björnholm B., Joergensen F.S., Youson J.H.,
  Schwartz T.W.;
RT "Primary structure and conformational analysis of peptide methionine-
  tyrosine, a peptide related to neuropeptide Y and peptide YY isolated
  from lamprey intestine.";
RL Eur. J. Biochem. 199;293-298(1991).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE NPV / PPV / PYY FAMILY.
DR PIR; S16943; S16943.
DR HSSP; P01303; IRON.
DR InterPro; IPR001955; Pancreatic_hormn.

```

```

DR Pfam; PF00159; hormone3; 1.
DR PRINTS; PR00278; PANC_HORMONE.
DR ProDom; PD001267; Pancreatic_hormn; 1.
DR SMART; SM00309; PAH; 1.
DR PROSITE; PS00265; PANCREATIC_HORMONE_1; 1.
DR PROSITE; PS50276; PANCREATIC_HORMONE_2; 1.
KW Hormone; Amidation.
FT MOD RES 36
SQ SEQUENCE 36 AA; 4206 MW; 76B5D50BED01B90A CRC64;

Query Match 44.8%; Score 26; DB 1; Length 36;
Best Local Similarity 55.6%; Pred. No. 65;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 LIRYLAIR 12
DB 17 LSKYLAIR 25

RESULT 4
CPRP_CANPG STANDARD; PRT; 38 AA.
AC P81033;
DT 15-JUL-1998 (Rel. 36; Created)
DT 15-JUL-1998 (Rel. 36; Last sequence update)
DT 15-DEC-1998 (Rel. 37; Last annotation update)
DE CHH precursor related peptide (CPRP).
OS Cancer pagurus (Rock crab).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Crustacea;
  Malacostraca; Eumalacostraca; Eucarida; Decapoda; Pleocyemata;
  Brachyura; Brachyura; Cancroidea; Cancridae; Cancer.
OX NCBI_TaxID=6755;
RN [1]
RP SEQUENCE.
RC TISSUE=Sinus gland;
RX MEDLINE=99025664; PubMed=9809792;
RA Chung J.S., Wilkenson M.C., Webster S.G.;
RT "Amino acid sequences of both isoforms of crustacean hyperglycemic
  RT hormone (CHH) and corresponding precursor-related peptide in Cancer
  pagurus.";
RL Regul. Pept. 77:17-24(1998).
CC -1- TISSUE SPECIFICITY: PRODUCED BY THE MEDULLA TERMINALIS X-ORGAN IN
  THE EYESTALKS AND TRANSPORTED TO THE SINUS GLAND WHERE IT IS
  STORED AND RELEASED.
KW Neuropeptide; Hormone.
SQ SEQUENCE 38 AA; 3969 MW; C979C87EB31ABB90 CRC64;

Query Match 43.1%; Score 25; DB 1; Length 38;
Best Local Similarity 60.0%; Pred. No. 1,1e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 RLIRYRLAI 11
DB 11 RLIRYRGL 20

RESULT 5
VPU_HVIA2 STANDARD; PRT; 38 AA.
AC P05549;
DT 01-NOV-1988 (Rel. 09; Created)
DT 01-NOV-1988 (Rel. 09; Last sequence update)
DT 01-NOV-1997 (Rel. 35; Last annotation update)
DE VPU protein (U ORF protein).
OS Human immunodeficiency virus type 1 (ARV2/SF2 isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11685;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85090453; PubMed=2578227;
RA Sanchez-Pescador R., Power M.D., Barr P.J., Steiner K.S.,
  Stempfen W.M., Brown-Shimer S.L., Gee W.W., Renard A., Randolph A.,

```


RA Levy J.A., Dina D., Luciw P.A.;
RT "Nucleotide sequence and expression of an AIDS-associated retrovirus
RL (ARV-2).";
RL Science 227:484-492(1985).
CC -!- FUNCTION: ACTS IN THE DEGRADATION OF CD4 IN THE ENDOPLASMIC
CC RETICULUM AND IN THE ENHANCEMENT OF VIRION RELEASE FROM THE PLASMA
CC MEMBRANE OF INFECTED CELLS.
CC -!- SUBCELLULAR LOCATION: Membrane-bound.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; K02007; AAB59881.1; --
DR HIV; K02007; VFU\$SF2.
DR InterPro; IPR002094; Vpu.
DR Pfam; PF00558; Vpu; 1.
KW Transmembrane; AIDS.
FT TRANSMEM 8 28 POTENTIAL.
SQ SEQUENCE 38 AA; 4390 MW; 786431F505BED6C7 CRC64;

Query Match 43.1%; Score 25; DB 1; Length 38;
Best Local Similarity 50.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 3 LLIRYLRLAIR 12
DB 26 VLIEYRKILR 35.

RESULT 6
ID Y382_TREPA STANDARD; PRT; 40 AA.
AC O83397;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein TP0382.
GN Treponema pallidum.
OS Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
OC NCBI_TaxID=160;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nichols;
RX Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
RA Dodson R., Winn M., Hickey E.K., Clayton R., Ketchum K.A.,
RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T.,
RA McDonald L., Artach P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA Venter J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis
RT spirochete";
RL Science 281:375-388(1998).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; A2001217; AAC65383.1; --
DR TIGR; TP0382; --
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 40 AA; 4104 MW; 1B74B6D9597CCC90 CRC64;

Query Match 43.1%; Score 25; DB 1; Length 40;
Best Local Similarity 44.4%; Pred. No. 1.1e+02;
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
QY 4 LLIRYLRLAIR 12
DB 19 LVNRLRLTMQ 27

RESULT 7
ID Y718_ARCFU STANDARD; PRT; 46 AA.
AC O29540;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein AF0718.
GN AF0718.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC Archaeoglobaceae; Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Winn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.P., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artach P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus";
RL Nature 390:364-370(1997).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB001055; AAB90531.1; --
DR TIGR; AF0718; --
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 46 AA; 5271 MW; 3AC6F5E54D58F9C5 CRC64;

Query Match 43.1%; Score 25; DB 1; Length 46;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 5 IRYRL 9
DB 42 IRYRL 46

RESULT 8
ID VSH_MUMPB STANDARD; PRT; 57 AA.
AC P28082;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Small hydrophobic protein.
GN SH.

```

OS Mumps virus (strain Belfast)..
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Rubulavirinae.
OX NCBI_TaxID=11166;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93168036; PubMed=8435047;
RA Yeo R.P., Afzal M.A., Forsey T., Rima B.K.;
RT "Identification of a new mumps virus lineage by nucleotide sequence
RT analysis of the SH gene of ten different strains.";
RL Arch. Virol. 128:371-377(1993).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X63709; CAA45242.1; -.
DR PIR; S19866; SHNZBP.
DR InterPro; IPR001477; SH.
DR Pfam; PF01445; SH; 1.
DR ProDom; PD001504; SH; 1.
KM Transmembrane.
FT TRANSMEM 8
SQ SEQUENCE 57 AA; 6814 MW; B8F2F917B4A8EF3C CRC64;

Query Match 42.2%; Score 24.5; DB 1; Length 57;
Best Local Similarity 28.6%; Pred. No. 2.1e+02;
Matches 6; Conservative 4; Mismatches 2; Indels 9; Gaps 1;

Qy 1 YRL-----RYRLAIR 12
   |||:|
   19 YRIITLYVWVSTIFYKYAVR 39
Db

RESULT 9
HSP1_PONPY STANDARD; PRT; 50 AA.
AC P35310;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Sperm protamine P1 (Cysteine-rich protamine).
GN PRM1.
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Butcheria; Primates; Catarrhini; Hominoidea; Pongo.
OX NCBI_TaxID=9600;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9411943; PubMed=8308910;
RA Relief J.D., Winfein R.J., Dixon G.H., Adroer R., Queralt R.,
RA Ballabriga J., Oliva R.;
RT "Evolution of protamine P1 genes in primates.";
RL J. Mol. Evol. 37:426-434(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=2012344; PubMed=10659848;
RA Wyckoff G.J., Wang W., Wu C.-I.;
RT "Rapid evolution of male reproductive genes in the descent of man.";
RL Nature 403:304-309(2000).
CC -1- FUNCTION: PROTAGMINES SUBSTITUTE FOR HISTONES IN THE CHROMATIN OF
CC Sperm DURING THE HAPLOID PHASE OF SPERMATOGENESIS. THEY COMPACT
CC Sperm DNA INTO A HIGHLY CONDENSED, STABLE AND INACTIVE COMPLEX.
CC -1- SUBUNIT: CROSS-LINKED BY INTERCHAIN DISULFIDE BONDS AROUND THE
CC DNA-HELIX (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: TESTIS.
CC -1- SIMILARITY: BELONGS TO THE PROTAGMINE P1 FAMILY.
CC -----

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U14589; AAA36946.1; -.
DR EMBL; AF215710; AAF34623.1; -.
DR InterPro; IPR000221; Protamine P1.
DR Pfam; PF00260; protamine P1; 1.
DR PROSITE; PS00048; PROTAGMINE P1; 1.
KM Chromosomal protein, Nucleosome core, Spermatogenesis; DNA-binding;
KM Testis; DNA condensation; Nuclear protein.
FT INIT MET 0
FT BY SIMILARITY.
SQ SEQUENCE 50 AA; 6660 MW; D9935BF411E5BE78 CRC64;

Query Match 41.4%; Score 24; DB 1; Length 50;
Best Local Similarity 71.4%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 6 RYRLAIR 12
   |||||
   Db 42 RYRLRCR 48

RESULT 10
LT01_HORVU STANDARD; PRT; 54 AA.
AC Q42509;
DT 15-JUN-1998 (Rel. 36, Created)
DT 15-JUN-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Low-temperature induced protein lt101.1 (Blt101) (Salt-stress induced
DE hydrophobic peptide ES13).
GN LT101.1 OR Lt101 OR ES13.
OS Hordeum vulgare (Barley), and
OS Lophopyrum elongatum (Tall wheatgrass) (Argropyrum elongatum).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticeae; Hordeum.
OX NCBI_TaxID=4513, 4588;
RN [1]
RP SEQUENCE FROM N.A.
RX SPECIES=H.vulgare; STRAIN=cv. Igri; TISSUE=Meristem;
RC MEDLINE=94072734; PubMed=8251639;
RA Goddard N.J., Dunn A.M., Zhang L., White A.J., Jack P.L., Hughes M.A.;
RT "Molecular analysis and spatial expression pattern of a low-
RT temperature-specific barley gene, blt101.";
RL Plant Mol. Biol. 23:871-879(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX SPECIES=H.vulgare; STRAIN=cv. Igri;
RC Brown A.P.C., Dunn A.M., Goddard N.J., Hughes M.A.;
RT "Identification of a novel low-temperature-response element in the
RT promoter of the barley gene lt101.1.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX SPECIES=L.elongatum; TISSUE=Root;
RC MEDLINE=94211914; PubMed=8159795;
RA Gulick P.J., Shen W., An H.;
RT "ES13, a stress-induced gene from Lophopyrum elongatum.";
RL Plant Physiol. 104:799-800(1994).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- INDUCTION: BY COLD SHOCK AND SALT STRESS.
CC -1- SIMILARITY: BELONGS TO THE UPF0057 (PM3) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; Z25537; CAA80984.1; -;
 DR EMBL; AJ310994; CAC37081.1; -;
 DR EMBL; U00966; AAA21847.1; -;
 DR InterPro; IPR000612; UPF0057.
 DR Pfam; PF01679; UPF0057; 1.
 DR PROSITE; PS01309; UPF0057; 1.
 KW Transmembrane.
 FT TRANSMEM 2 22 POTENTIAL.
 FT TRANSMEM 34 54 POTENTIAL.
 SQ SEQUENCE 54 AA; 5936 MW; 52521C39468B8BF2 CRC64;

Query Match 41.4%; Score 24; DB 1; Length 54;
 Best Local Similarity 33.4%; Pred. No. 2.5e+02;
 Matches 3; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 3 LLIRYRLAI 11
 : : : : :
 DB 20 VFLRYKLG 28

RESULT 11
 ID L702_HORVU STANDARD; PRT; 54 AA.
 AC Q9ARD5;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Low-temperature induced protein lt101.2.
 GN Lt101.2.
 OS Hordeum vulgare (Barley).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
 OC Triticeae; Hordeum.
 OX NCBI_TaxID=4513;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Igri.
 RX MEDLINE=21533241; PubMed=11678282;
 RA Brown A.P.C., Dunn M.A., Goddard N.J., Hughes M.A.;
 RT "Identification of a novel low-temperature-response element in the
 RT promoter of the barley (Hordeum vulgare L) gene blt101.1.";
 RL Planta 213:770-780(2001).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -!- INDUCTION: BY COLD SHOCK AND SALT STRESS.
 CC -!- SIMILARITY: BELONGS TO THE UPF0057 (PMP3) FAMILY.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; AJ310995; CAC37082.1; -;
 DR InterPro; IPR000612; UPF0057.
 DR Pfam; PF01679; UPF0057; 1.
 DR PROSITE; PS01309; UPF0057; 1.
 KW Transmembrane.
 FT TRANSMEM 2 22 POTENTIAL.
 FT TRANSMEM 34 54 POTENTIAL.
 SQ SEQUENCE 54 AA; 5913 MW; 7C2FC7482244A053 CRC64;

Query Match 41.4%; Score 24; DB 1; Length 54;
 Best Local Similarity 44.4%; Pred. No. 2.5e+02;
 Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 LLIRYRLAI 11
 : : : : :
 DB 20 VFLRYKLG 28

DB 20 VFLRYGLAV 28

RESULT 12
 ID LCRT_YERPS STANDARD; PRT; 56 AA.
 AC Q00932;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 01-JUL-1993 (Rel. 26, Last annotation update)
 DE Low calcium response locus protein T.
 GN LCRT.
 OS Yersinia pseudotuberculosis.
 OG Plasmid pIB1.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Yersinia.
 OX NCBI_TaxID=633;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=YPIII;
 RX MEDLINE=92250432; PubMed=1577700;
 RA Rimpilainen M., Forsberg A., Wolf-Watz H.;
 RT "A novel protein, LcrQ, involved in the low-calcium response of
 RT Yersinia pseudotuberculosis shows extensive homology to YopH.";
 RL J. Bacteriol. 174:3355-3363(1992).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; M83986; AAA227656.1; -;
 DR InterPro; IPR001230; Prenyl_site.
 KW Plasmid.
 SQ SEQUENCE 56 AA; 6760 MW; B6C9AD6C40A16593 CRC64;

Query Match 41.4%; Score 24; DB 1; Length 56;
 Best Local Similarity 62.5%; Pred. No. 2.6e+02;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 YRLIRYR 8
 : : : : :
 DB 32 YRLRYRK 39

RESULT 13
 ID Y195_BPT7 STANDARD; PRT; 49 AA.
 AC P03604;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 01-MAR-1989 (Rel. 10, Last annotation update)
 DE Hypothetical gene 19.5 protein.
 GN 19.5.
 OS Bacteriophage T7.
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;
 OC T7-like viruses.
 OX NCBI_TaxID=10760;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=83241725; PubMed=6864790;
 RA Dunn J.J., Studier F.W.;
 RT "Complete nucleotide sequence of bacteriophage T7 DNA and the
 RT locations of T7 genetic elements.";
 RL J. Mol. Biol. 166:477-535(1983).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: V01146; CAA24443.1; -.
DR PIR: A04429; O9BPE7.
DR PIR: S42341; S42341.
DR Hypothetical protein.
KW Hypothetical protein.
SQ SEQUENCE 49 AA; 5434 MW; FB1C9240BA465342 CRC64;

Query Match 40.5%; Score 23.5; DB 1; Length 49;
Best Local Similarity 40.0%; Pred. No. 2.9e+02;
Matches 6; Conservative 4; Mismatches 2; Indels 3; Gaps 1;

Oy 1 YRLIRYRL 9
   |||::|:|
Db 2 FRLNLRLRHYR 16

RESULT 14
Y573 TREPA STANDARD; PRT; 30 AA.
AC O83583;
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein TP0573.
GN TP0573.
OS Treponema pallidum.
OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxId=160;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nichols;
RX MEDLINE=98332770; PubMed=9665876;
RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
RA Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,
RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Utechtack T.,
RA McDonald L., Artlich P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA Venter J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis
RT spirochete."
RL Science 281:375-388(1998).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE001232; AAC65550.1; -.
DR TIGR: TP0573; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 30 AA; 3380 MW; COB2640A4B30E33C CRC64;

Query Match 39.7%; Score 23; DB 1; Length 30;
Best Local Similarity 33.3%; Pred. No. 2.1e+02;
Matches 3; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Oy 1 YRLIRYRL 9
   |||::|:|
Db 4 FSLVAVYKI 12

RESULT 15
PY AMICA STANDARD; PRT; 36 AA.
ID PY AMICA
AC P29705;
DT 01-DEC-1992 (Rel. 24, Created)

```

```

DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Peptide YX-1like (PYX).
DE PYX.
GN PYX.
OS Amla calva (Bowfin).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Amliformes; Amlidae; Amla.
OX NCBI_TaxId=7924;
RN [1]
RP SEQUENCE.
RC TISSUE=Pancreas;
RX MEDLINE=91296574; PubMed=2067973;
RA Conlon J.M., Bjennning C., Moon T.W., Youson J.H., Thim L.;
RT "Neuropeptide Y-related peptides from the pancreas of a teleostean
RL Peptides 12:221-226(1991).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE NPY / PPY / PYY FAMILY.
DR HSP; P01303; IRON.
DR InterPro; IPR001955; Pancreatic_hormn.
DR Pfam; PF00159; hormones; 1.
DR PRINTS; PR00278; PANCHORMONE.
DR ProDom; PD001267; Pancreatic_horma; 1.
DR SMART; SMO0309; PAH; 1.
DR PROSITE; PS00265; PANCREATIC_HORMONE_1; 1.
DR PROSITE; PS50276; PANCHREATIC_HORMONE_2; 1.
KW Hormone; Amidation.
FT MOD_RES 36
FT MOD_RES 36
SQ SEQUENCE 36 AA; 4333 MW; 56B46F3C08666671 CRC64;

Query Match 39.7%; Score 23; DB 1; Length 36;
Best Local Similarity 55.6%; Pred. No. 2.6e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Oy 4 LIRYRLAIR 12
   |||::|:|
Db 17 LARYTLAIR 25

Search completed: December 19, 2002, 16:19:46
Job time : 6.47059 secs

```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 19, 2002, 16:16:02 ; Search time 10.4118 Seconds
(without alignments)
110.799 Million cell updates/sec

Title: US-08-653-294C-42
Perfect score: 58
Sequence: 1 YRLIRYRLAIR 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 15911

Minimum DB seq length: 0
Maximum DB seq length: 60

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30	51.7	43	AH0876	conserved hypothet
2	28.5	49.1	44	IMBP8	hypothetical immun
3	28	48.3	25	S35926	T-cell receptor ga
4	28	48.3	39	G90716	probable RNA (impo
5	28	48.3	39	G64801	hypothetical prote
6	27	46.6	53	B83953	hypothetical prote
7	27	46.6	54	F72675	hypothetical prote
8	27	46.6	57	JQ2368	small hydrophobic
9	27	46.6	57	AL2525	transposase asr738
10	26	44.8	26	S04376	hypothetical prote
11	26	44.8	36	S16943	neuropeptide Y - s
12	26	44.8	50	B70239	hypothetical prote
13	26	44.8	55	T09464	H4-transporting tw
14	25	43.1	29	B81006	hypothetical prote
15	25	43.1	40	A70237	hypothetical prote
16	25	43.1	40	F73330	hypothetical prote
17	25	43.1	46	H69339	hypothetical prote
18	25	43.1	53	C82714	S08 ribosomal prot
19	24.5	42.2	57	SHNZBF	small hydrophobic
20	24	41.4	23	T50545	reductase (impor
21	24	41.4	30	A05012	hypothetical prote
22	24	41.4	30	PD0006	cysteine synthase
23	24	41.4	33	S42781	relaxin - oranguta
24	24	41.4	33	B82184	hypothetical prote
25	24	41.4	51	S72280	hypothetical prote
26	24	41.4	54	S40406	bit101 protein - b
27	24	41.4	56	B89903	hypothetical prote
28	24	41.4	57	F82816	hypothetical prote
29	24	41.4	58	A86702	hypothetical prote

hypothetical prote
integral membrane
hypothetical prote
gene 19.5 protein
relaxin - baboon (
hypothetical prote
hypothetical prote
relaxin - hamadrya
hypothetical prote
probable succinate
pancreatic hormone
hypothetical prote
hypothetical prote
ribosomal protein
gene 39.2 protein
hypothetical prote

ALIGNMENTS

RESULT 1

AH0876

conserved hypothetical protein STY3241 [imported] - Salmonella enterica subsp. enterica
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: this species has also been called Salmonella typhi
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 09-Nov-2001
C:Accession: AH0876
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Conington, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica sero
A:Reference number: AB0502; PMID:11677608
A:Accession: AH0876
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-43 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD02912.1; PID:g16504165; GSPDB:GN00176
C:Genetics:
A:Gene: STY3241

Query Match 51.7%; Score 30; DB 2; Length 43;
Best Local Similarity 63.6%; Pred. No. 49;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 YRLIRYRLAI 11
| | | | |
DB 22 YGLSLRLAI 32

RESULT 2

IMBP8

hypothetical immunity region protein 8 - Bacillus phage phi-105
C:Species: Bacillus phage phi-105
A:Note: host Bacillus subtilis
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
C:Accession: G24521
R:Cully, D.F.; Garro, A.J.
Gene 38, 153-164, 1985
A:Title: Nucleotide sequence of the immunity region of Bacillus subtilis bacteriophage i
A:Reference number: A91535; MUID:86056972; PMID:3934047
A:Accession: G24521
A:Molecule type: DNA
A:Residues: 1-44 <CUL>
A:Cross-references: GB:M1920; NID:g215477
C:Comment: This is the hypothetical translation of a sequence that was not reported as

Query Match 49.1%; Score 28.5; DB 4; Length 44;
Best Local Similarity 63.6%; Pred. No. 99;
Matches 7; Conservative 3; Mismatches 0; Indels 1; Gaps 1;

OY 3 LLIRY-RLAIR 12
 Db 9 LLIRYRLALR 19

RESULT 3

S35926
 T-cell receptor gamma chain - human (fragment)
 C:Species: Homo sapiens (man)
 C:Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 30-Jun-2001
 C:Accession: S35926
 R:Machoudakis, G.; Platsoucas, C.D.
 submitted to the EMBL Data Library, May 1993
 A:Description: An alternative splicing between V-gamma, J-gamma2.3 and C-gamma2 gene seg
 A:Reference number: S32764
 A:Accession: S35926
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-25 <MAT>
 A:Cross-references: EMBL:222688
 C:Keywords: T-cell receptor

Query Match 48.3%; Score 28; DB 2; Length 25;
 Best Local Similarity 54.5%; Pred. No. 72;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 YLLIRYRLAI 11
 Db 7 YLLIRNSLAV 17

RESULT 4

G90716
 Probable RNA [imported] - Escherichia coli (strain O157:H7, substrain RMD 0509952)
 C:Species: Escherichia coli
 C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
 C:Accession: G90716
 R:Hayashi, T.; Makino, K.; Onishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
 gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.
 DNA Res. 8, 11-22, 2001
 A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc
 A:Reference number: A99629; MUID:21156231; PMID:11258796
 A:Accession: G90716
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-39 <HAY>
 A:Cross-references: GB:BA000007; PIDN:BA834126.1; PID:G13360161; GSPDB:GN00154
 A:Experimental source: strain O157:H7, substrain RMD 0509952
 C:Genetics:
 A:Gene: EC90703

Query Match 48.3%; Score 28; DB 2; Length 39;
 Best Local Similarity 54.5%; Pred. No. 1.1e+02;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 YLLIRYRLAI 11
 Db 21 YOKPVPYRLAI 31

RESULT 5

G64801
 Hypothetical protein b0669 - Escherichia coli (strain K-12)
 C:Species: Escherichia coli
 C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002
 C:Accession: G64801
 R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
 A.; Rose, D.J.; Mau, B.; Shao, Y.
 Science 277, 1453-1462, 1997
 A:Title: The complete genome sequence of Escherichia coli K-12.
 A:Reference number: A64720; MUID:97426617; PMID:9278503
 A:Accession: G64801
 A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA
 A:Residues: 1-39 <BLAT>
 A:Cross-references: GB:AE000170; GB:U00096; NID:G1786875; PIDN:AACT3766.1; PID:G1786866;
 A:Experimental source: strain K-12, substrain MG1655

Query Match 48.3%; Score 28; DB 2; Length 39;
 Best Local Similarity 54.5%; Pred. No. 1.1e+02;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 YLLIRYRLAI 11
 Db 21 YOKPVPYRLAI 31

RESULT 6

B83853
 Hypothetical protein BH1626 [imported] - Bacillus halodurans (strain C-125)
 C:Species: Bacillus halodurans
 C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
 C:Accession: B83853
 R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira
 Nucleic Acids Res. 28, 4317-4331, 2000
 A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
 A:Reference number: A83650; MUID:20512582; PMID:11058132
 A:Accession: B83853
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-53 <STO>
 A:Cross-references: GB:AP001512; GB:BA000004; NID:G10174030; PIDN:BA805345.1; GSPDB:GN00
 A:Experimental source: strain C-125
 C:Genetics:
 A:Gene: BH1626

Query Match 46.6%; Score 27; DB 2; Length 53;
 Best Local Similarity 71.4%; Pred. No. 2.3e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 6 RYRLAIR 12
 Db 5 RYRVIR 11

RESULT 7

F72675
 Hypothetical protein APES036 - Aeropyrum pernix (strain K1)
 C:Species: Aeropyrum pernix
 C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
 C:Accession: F72675
 R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah
 awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Ki
 DNA Res. 6, 83-101, 1999
 A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
 A:Reference number: A72450; MUID:99310339; PMID:10382966
 A:Accession: F72675
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-54 <KAW>
 A:Cross-references: DDBJ:AP000060; NID:G5104188; PIDN:BA479806.1; PID:G1043592; PID:G510
 A:Experimental source: strain K1
 C:Genetics:
 A:Gene: APES036

Query Match 46.6%; Score 27; DB 2; Length 54;
 Best Local Similarity 71.4%; Pred. No. 2.4e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 3 LLIRYRL 9
 Db 18 LLIRRL 24

RESULT 8

J02368

small hydrophobic protein - mumps virus

C:Species: mumps virus
C>Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 07-May-1999
C:Accession: JQ2368
R:Afzal, M.A.; Pickford, A.R.; Forsey, T.; Heath, A.B.; Minor, P.D.
J. Gen. Virol. 74, 917-920, 1993
A:Title: The Jeryl Lynn vaccine strain of mumps virus is a mixture of two distinct isolates
A:Reference number: JQ2366; MUID:93260409; PMID:8492099
A:Accession: JQ2368
A:Molecule type: mRNA
A:Residues: 1-57 <APZ>
C:Genetics:
A:Gene: SH
C:Superfamily: mumps virus small hydrophobic protein
C:Keywords: membrane protein

Query Match 46.6%; Score 27; DB 2; Length 57;
Best Local Similarity 50.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 LLIRYRLAIR 12
| | | | |
DB 30 LTINYKTAVR 39

RESULT 9
A12525
transposase asr7385 [imported] - Nostoc sp. (strain PCC 7120) plasmid pCC7120alpha
C:Species: Nostoc sp.
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
C:Accession: A12525
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kunitz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: A12525
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-57 <KUR>
A:Cross-references: GB:BA000020; PIDN:BA077143.1; PID:G17134584; GSPDB:GN00180
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: asr7385
A:Genome: plasmid

Query Match 46.6%; Score 27; DB 2; Length 57;
Best Local Similarity 60.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 RLIRYRLAIR 11
| | | | |
DB 38 RLIRYRLAIR 47

RESULT 10
S04376
hypothetical protein - Klebsiella pneumoniae (fragment)
C:Species: Klebsiella pneumoniae
C>Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 26-Aug-1999
C:Accession: S04376
R:Holtel, A.; Merrick, M.
Mol. Gen. Genet. 215, 134-138, 1988
A:Title: Identification of the Klebsiella pneumoniae glbB gene: nucleotide sequence of
A:Reference number: S04376; MUID:89201333; PMID:2907369
A:Accession: S04376
A:Molecule type: DNA
A:Residues: 1-26 <HOL>
A:Cross-references: EMBL:X14012; NID:G43804; PIDN:CAA32176.1; PID:G43805
C:Superfamily: nitrogen assimilation regulatory protein ntrC; response regulator homolog
Query Match 44.8%; Score 26; DB 2; Length 26;

Best Local Similarity 55.6%; Pred. No. 1.8e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 YRLIRYRL 9
| | | | |
DB 11 YKLLSRHEL 19

RESULT 11
S16943
neuropeptide Y - sea lamprey
C:Species: Petromyzon marinus (sea lamprey)
C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 31-Oct-1997
C:Accession: S16943
R:Conlon, J.M.; Bjornholm, B.; Jorgensen, P.S.; Youson, J.H.; Schwartz, T.W.
Eur. J. Biochem. 199, 293-298, 1991
A:Title: Primary structure and conformational analysis of peptide methionine-tyrosine,
A:Reference number: S16943; MUID:91301137; PMID:2070789
A:Accession: S16943
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-36 <CON>
C:Superfamily: pancreatic hormone

Query Match 44.8%; Score 26; DB 2; Length 36;
Best Local Similarity 55.6%; Pred. No. 2.5e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 LRIRYRLAIR 12
| | | | |
DB 17 LSKYMLAVR 25

RESULT 12
B70239
hypothetical protein BBH36 - Lyme disease spirochete plasmid H/lp28-3
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C>Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 08-Oct-1999
C:Accession: B70239
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Whitson, D.; Peterson, J.; Karlavag, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt, B.; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A:Authors: Smith, H.O.; Venter, J.C.
A:Title: Genomic sequence of a Lyme disease spirochete,
A:Reference number: A70100; MUID:98065943; PMID:9403685
A:Accession: B70239
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-50 <KLE>
A:Cross-references: GB:AE000784; NID:G2690041; PIDN:AAC66017.1; PID:G2690073; TIGR:BBH36
A:Experimental source: strain B31
C:Genetics:
A:Genome: plasmid

Query Match 44.8%; Score 26; DB 2; Length 50;
Best Local Similarity 54.5%; Pred. No. 3.4e+02;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 YRLIRYRLAIR 11
| | | | |
DB 36 YLLIIFRPSI 46

RESULT 13
T09464
H+-transporting two-sector ATPase (EC 3.6.3.14) chain 8 - brown alga (Pylaiella littoralis)
C:Species: Pylaiella littoralis
C>Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 03-Jun-2002
C:Accession: T09464
R:Aouvois, S.; Oudot, M.; Fontaine, J.; Kloareg, B.; Goer, S.L.
J. Mol. Biol. 277, 1047-1057, 1998
A:Title: Witnessing the evolution of transcription in mitochondria: The mitochondrial ge

A:Reference number: Z16681; MUID:98239704; PMID:9571021

A:Accession: T09464

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-55 <ROU>

A:Cross-references: EMBL:AF034976; NID:93243103; PID:93243105

A:Experimental source: strain Roscoff

C:Genetics:

A:Gene: atp8

A:Genome: mitochondrion

C:Keywords: ATP biosynthesis; hydrolase; mitochondrion; oxidative phosphorylation

Query Match

44.8%; Score 26; DB 2; Length 55;

Best Local Similarity 36.4%; Pred. No. 3.7e+02;

Matches 4; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 YRLIRYRLAI 11

DB 24 YMTVRYMLPV 34

RESULT 14

B81006

hypothetical protein NMB2097 [imported] - Neisseria meningitidis (strain MC58 serogroup

C:Species: Neisseria meningitidis

C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001

C:Accession: B81006

R:Retellin, H.; Saunders, N.J.; Heideberg, J.; Jeffries, A.C.; Nelson, K.E.; Eissen, J.A.

Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;

ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masiarant, V.; Pizza, M.

Science 287, 1809-1815, 2000

A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve

A>Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.

A:Reference number: A81000; MUID:2015755; PMID:10710307

A:Accession: B81006

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-29 <TET>

A:Cross-references: GB:AE002559; GB:AE002098; NID:97227357; PIDN:AAF42414.1; PID:9722736

A:Experimental source: serogroup B, strain MC58

C:Genetics:

A:Gene: NMB2097

Query Match

43.1%; Score 25; DB 2; Length 29;

Best Local Similarity 50.0%; Pred. No. 3.2e+02;

Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 YRLIRYR 8

DB 20 FSLLRFR 27

RESULT 15

A70237

hypothetical protein BBH16 - Lyme disease spirochete plasmid H/1p28-3

C:Species: Borrelia burgdorferi (Lyme disease spirochete)

C>Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 08-Oct-1999

C:Accession: A70237

R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White

son, D.; Peterson, J.; Kervatage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,

; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.

Nature 390, 580-586, 1997

A:Authors: Smith, H.O.; Venter, J.C.

A>Title: Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.

A:Reference number: A70100; MUID:98065943; PMID:9403685

A:Accession: A70237

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-40 <KLB>

A:Cross-references: GB:AE000784; NID:92690041; PIDN:AA66020.1; PID:92690076; TIGR:BBH16

A:Experimental source: strain B31

C:Genetics:

A:Genome: plasmid

Query Match 43.1%; Score 25; DB 2; Length 40;

Best Local Similarity 60.0%; Pred. No. 4.3e+02;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 RLLIRYRLAI 11

DB 20 QVLRKRYRLAI 29

Search completed: December 19, 2002, 16:23:00
Job time : 11.418 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 19, 2002, 16:02:26 ; Search time 28.5882 Seconds
(without alignments)
55.932 Million cell updates/sec

Title: US-08-653-294C-41
Perfect score: 58
Sequence: 1 RYALRYRILLRY 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 473727

Minimum DB seq length: 0
Maximum DB seq length: 60

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	34	58.6	49	AAU22858	Human prostate cancer
2	34	58.6	49	AAU22858	Human prostate cancer
3	33	56.9	38	AAU22858	Human prostate cancer
4	33	56.9	38	AAU22858	Human prostate cancer
5	33	56.9	38	AAU22858	Human prostate cancer
6	33	56.9	38	AAU22858	Human prostate cancer
7	33	56.9	38	AAU22858	Human prostate cancer
8	32	55.2	39	AAU22858	Human prostate cancer
9	32	55.2	54	AAU22858	Human prostate cancer
10	32	55.2	54	AAU22858	Human prostate cancer

11	32	55.2	54	22	AAU22858	Human prostate cancer antigen, Seq ID No 377.
12	32	55.2	54	22	AAU22858	Human prostate cancer antigen, Seq ID No 377.
13	31.5	54.3	52	16	AAU22858	Human prostate cancer antigen, Seq ID No 377.
14	31.5	54.3	52	16	AAU22858	Human prostate cancer antigen, Seq ID No 377.
15	31.5	54.3	52	16	AAU22858	Human prostate cancer antigen, Seq ID No 377.
16	31	53.4	53	20	AAU22858	Human prostate cancer antigen, Seq ID No 377.
17	30	51.7	20	18	AAU22858	Human prostate cancer antigen, Seq ID No 377.
18	30	51.7	21	22	AAU22858	Human prostate cancer antigen, Seq ID No 377.
19	30	51.7	34	23	AAU22858	Human prostate cancer antigen, Seq ID No 377.
20	30	51.7	40	22	AAU22858	Human prostate cancer antigen, Seq ID No 377.
21	30	51.7	57	22	AAU22858	Human prostate cancer antigen, Seq ID No 377.
22	30	51.7	57	22	AAU22858	Human prostate cancer antigen, Seq ID No 377.
23	30	51.7	57	22	AAU22858	Human prostate cancer antigen, Seq ID No 377.
24	30	51.7	57	22	AAU22858	Human prostate cancer antigen, Seq ID No 377.
25	30	51.7	57	22	AAU22858	Human prostate cancer antigen, Seq ID No 377.
26	30	51.7	57	22	AAU22858	Human prostate cancer antigen, Seq ID No 377.
27	30	51.7	57	22	AAU22858	Human prostate cancer antigen, Seq ID No 377.
28	30	51.7	57	22	AAU22858	Human prostate cancer antigen, Seq ID No 377.
29	30	51.7	57	22	AAU22858	Human prostate cancer antigen, Seq ID No 377.
30	30	51.7	57	22	AAU22858	Human prostate cancer antigen, Seq ID No 377.
31	29	50.0	6	19	AAU22858	Human prostate cancer antigen, Seq ID No 377.
32	29	50.0	6	19	AAU22858	Human prostate cancer antigen, Seq ID No 377.
33	29	50.0	6	19	AAU22858	Human prostate cancer antigen, Seq ID No 377.
34	29	50.0	6	19	AAU22858	Human prostate cancer antigen, Seq ID No 377.
35	29	50.0	9	22	AAU22858	Human prostate cancer antigen, Seq ID No 377.
36	29	50.0	9	22	AAU22858	Human prostate cancer antigen, Seq ID No 377.
37	29	50.0	10	14	AAU22858	Human prostate cancer antigen, Seq ID No 377.
38	29	50.0	10	16	AAU22858	Human prostate cancer antigen, Seq ID No 377.
39	29	50.0	10	16	AAU22858	Human prostate cancer antigen, Seq ID No 377.
40	29	50.0	10	16	AAU22858	Human prostate cancer antigen, Seq ID No 377.
41	29	50.0	10	16	AAU22858	Human prostate cancer antigen, Seq ID No 377.
42	29	50.0	10	16	AAU22858	Human prostate cancer antigen, Seq ID No 377.
43	29	50.0	10	16	AAU22858	Human prostate cancer antigen, Seq ID No 377.
44	29	50.0	10	17	AAU22858	Human prostate cancer antigen, Seq ID No 377.
45	29	50.0	10	17	AAU22858	Human prostate cancer antigen, Seq ID No 377.

ALIGNMENTS

RESULT 1
AAU22858
ID AAU22858 standard; Protein; 49 AA.
XX AC AAU22858;
XX AC
DT 17-DEC-2001 (first entry)
XX Human prostate cancer antigen, Seq ID No 377.
DE Human; prostate cancer antigen; cytostatic; uropathic; diagnostic;
KW reproductive system; chromosomal marker; forensic; urinary disorder;
KW chronic nephritis; blood-related disorder; thrombosis.
XX Homo sapiens.
XX WO200155316-A2.
XX PD
XX 02-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US01328.
XX PR 31-JAN-2000; 2000US-0179065.
XX PR 04-FEB-2000; 2000US-0180628.
XX PR 24-FEB-2000; 2000US-0184664.
XX PR 02-MAR-2000; 2000US-0186350.
XX PR 16-MAR-2000; 2000US-0189874.
XX PR 17-MAR-2000; 2000US-0190076.
XX PR 18-APR-2000; 2000US-0198123.
XX PR 19-MAY-2000; 2000US-0205515.
XX PR 07-JUN-2000; 2000US-0209467.
XX PR 28-JUN-2000; 2000US-0214886.
XX PR 30-JUN-2000; 2000US-0215135.

Peptide #7937 enco
Human peptide enco
HLA-B2702 84-79-84
Peptide B2702.84-7
Immunomodulating d
Rice E2F protein f
Tachyegrin and/or
Human interleukin-
Human ORF100 prote
Human immune/haema
Novel human connec
Peptide #3135 enco
Peptide #3157 enco
Protein #3079 enco
Human brain expres
Human bone marrow
Peptide #3101 enco
Peptide #3188 enco
Peptide #3065 enco
Human peptide enco
Peptide #3 used in
Peptide #1 used in
Immunomodulatory p
Immunosuppressive
Human Class I HLA-
Peptide fragment o
Alpha1-helix of HL
HLA-B2702 75-84(D)
HLA-B2702 CTL modu
HLA-B2702 CTL modu
HLA-B2702 CTL modu
T-cell modulating
T-cell modulating
Cytomodulating lip

CC for testing and detection e.g. as a chromosomal marker and in forensics.
CC (I) and the anti-(II) antibody can be used in testing and detection in
CC immunoassays. AAU22702-AAU22913 represent the human prostate cancer
CC antigen amino acid sequences, and related amino acid sequences of the
CC invention. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at: ftp.wipo.int/pub/published_pt_sequences.
XX
SQ Sequence 49 AA;

Query Match 58.6%; Score 34; DB 22; Length 49;
Best Local Similarity 58.3%; Pred. No. 33;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 RIALRYRILLRY 12
: ||| | : || |
Db 14 QIALSYNLLLY 25

RESULT 2

AAM96170
ID AAM96170 standard; Protein; 49 AA.

XX AC
XX AAM96170;

DT 21-NOV-2001 (first entry)

XX DE Human reproductive system related antigen SEQ ID NO: 4928.

XX KW Human; reproductive system related antigen; reproductive system disorder;
XX KM cancer; gene therapy.

XX OS Homo sapiens.

XX PN WO200155320-A2.

XX PD 02-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US01339.

XX 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.

PR 24-FEB-2000; 2000US-0184664.

PR 02-MAR-2000; 2000US-0186350.

PR 16-MAR-2000; 2000US-0189874.

PR 17-MAR-2000; 2000US-0190076.

PR 18-APR-2000; 2000US-0198123.

PR 19-MAY-2000; 2000US-0205515.

PR 07-JUN-2000; 2000US-0209467.

PR 28-JUN-2000; 2000US-0214886.

PR 30-JUN-2000; 2000US-0215135.

PR 07-JUL-2000; 2000US-0216647.

PR 07-JUL-2000; 2000US-0216880.

PR 11-JUL-2000; 2000US-0217487.

PR 11-JUL-2000; 2000US-0217496.

PR 14-JUL-2000; 2000US-0218290.

PR 26-JUL-2000; 2000US-0220963.

PR 26-JUL-2000; 2000US-0220964.

PR 14-AUG-2000; 2000US-0224518.

PR 14-AUG-2000; 2000US-0224519.

PR 14-AUG-2000; 2000US-0225113.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225266.

PR 14-AUG-2000; 2000US-0225267.

PR 14-AUG-2000; 2000US-0225268.

PR 14-AUG-2000; 2000US-0225270.

PR 14-AUG-2000; 2000US-0225447.

PR 14-AUG-2000; 2000US-0225757.

PR 14-AUG-2000; 2000US-0225758.

PR 14-AUG-2000; 2000US-0225759.

PR 18-AUG-2000; 2000US-0226279.

PR 22-AUG-2000; 2000US-0226681.

PR 22-AUG-2000; 2000US-0226868.

PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.

PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX Rosen CA, Barash SC, Ruben SM;
 PI WPI; 2001-465570/50.
 DR N-PSDB; AAL02140.
 XX
 PT Isolated nucleic acid molecule encoding a reproductive system antigen
 PT is used in preventing, treating or ameliorating a medical condition -
 XX
 PS Claim 11; SEQ ID NO 4828; 1297bp + Sequence Listing; English.
 CC The present invention provides the protein and coding sequences of a
 CC number of human reproductive system related antigens. These can be used
 CC in the prevention and treatment of reproductive system disorders,
 CC including cancer. The present sequence is a protein of the invention.
 XX
 SQ Sequence 49 AA;
 Query Match 58.6%; Score 34; DB 22; Length 49;
 Best Local Similarity 58.3%; Pred. No. 33;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 1 RTALRRILLRY 12
 :|||:|:
 Db 14 QIALSYNLLLY 25
 RESULT 3
 ID AAY88733 standard; peptide; 38 AA.
 XX AAY88733;
 AC
 XX
 DT 23-MAY-2000 (first entry)
 DE Core polypeptide fragment T No. 88.
 XX
 XX Retrovirus; hybrid polypeptide; enhancer; gp41; envelope protein; HIV-1;
 KM HIV-2; SIV; gammacocci; half-life; growth factor; cytokine; viral;
 KM anti-fusogenic; differentiation factor; interleukin; interferon;
 KM colony stimulating factor; hormone; angiogenic factor.

XX
 OS Unidentified.
 XX
 XX WO9959615-A1.
 XX
 PD 25-NOV-1999.
 XX
 PE 20-MAY-1999; 99WO-US11219.
 XX
 PR 20-MAY-1998; 98US-0082279.
 XX
 PA (TRIM-) TRIMERIS INC.
 PI Barney S, Guthrie KI, Merutka G, Anwer MK, Lambert DM;
 DR WPI; 2000-136792/12.
 XX
 PT A new hybrid polypeptide with enhanced pharmacokinetic properties
 PT comprises enhancer sequence -
 XX
 PS Disclosure; Page 22; 124pp; English.
 XX
 CC The invention relates to hybrid polypeptides comprising enhancer peptide
 CC sequence linked to core polypeptides. The enhancer polypeptides are
 CC derived from various retroviral envelope (gp41) protein sequences,
 CC especially from HIV-1, HIV-2 and SIV. The enhancer peptides enhance the
 CC pharmacokinetic properties such as increasing the half-life of any core
 CC polypeptide that they are linked to. The core polypeptides are any
 CC polypeptide that may be introduced into a living system and that can
 CC function as a pharmacologically useful peptide for the treatment or
 CC prevention of a disease. The core polypeptides are bioactive peptides
 CC selected from a growth factor, cytokine, differentiation factor,
 CC interleukin, interferon, colony stimulating factor, hormone or
 CC angiogenic factor. The peptides of the invention can be used for
 CC inhibiting viral infection and can be used in anti-viral and
 CC anti-fusogenic treatments. Sequences AAY88651-Y90055 represent core
 CC polypeptide fragments that can be used in the invention. Some sequences
 CC among those indicated also comprise enhancer fragments at terminal ends
 CC and form hybrid polypeptides.
 XX
 SQ Sequence 38 AA;
 Query Match 56.9%; Score 33; DB 21; Length 38;
 Best Local Similarity 45.5%; Pred. No. 38;
 Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
 QY 2 IALRRYILLRY 12
 :|||:|:
 Db 24 LAVREKVLIRY 34
 RESULT 4
 ID ABB00092 standard; Peptide; 38 AA.
 XX ABB00092;
 AC
 XX
 DT 03-JAN-2002 (first entry)
 DE Viral DP178/107-like region peptide T88.
 XX
 XX Human immunodeficiency virus; HIV; respiratory syncytial virus; RSV;
 KM viruslike; heptad repeat region; transmembrane protein; gp41; HR1; HR2;
 KM infection.
 XX
 OS Viridiae.
 XX
 XX
 FH Key Location/Qualifiers
 FT Modified-site 1
 FT Modified-site /note= "N-terminal is substituted by Ac"
 FT Modified-site 38
 FT /note= "C-terminal amide"
 XX

PN WO200164013-A2.
 XX PD 07-SEP-2001.
 XX XX 07-FEB-2001; 2001WO-US03988.
 XX PR 29-FEB-2000; 2000US-0515965.
 XX PA (TRIM-) TRIMERIS INC.
 XX PI Antczak JB, Delmedico MK, Erickson JB, Lambert DM, Sista P;
 XX DR WPI; 2001-514829/56.
 XX PT Heptad repeat region peptide analogs useful for inhibiting virus/cells
 XX PT fusion, useful for treating HIV and Respiratory Syncytial Virus
 XX PT infection -
 XX XX Disclosure; Page 34; 587pp; English.
 XX CC The invention relates to isolated analogues of the heptad repeat region
 XX CC peptides DP178 and DP107. DP178 and DP107 correspond to amino acids
 XX CC 638-673 (heptad repeat region HR2) and 558-595 (heptad repeat region
 XX CC HR1) respectively, of HIV-1LAI transmembrane protein gp41. The HR1
 XX CC and HR2 regions of proteins interact non-covalently with each other
 XX CC and/or with peptides derived from them. This interaction is required for
 XX CC normal infectivity of viruses such as RSV and HIV. The heptad
 XX CC repeat region peptide analogues may be used to inhibit respiratory
 XX CC syncytial virus (RSV) infection in a cell. They may also be used to
 XX CC inhibit HIV infection. The present sequence is a peptide provided in
 XX CC the specification.
 XX SQ Sequence 38 AA;
 XX Query Match 56.9%; Score 33; DB 22; Length 38;
 XX Best Local Similarity 45.5%; Pred. No. 38;
 XX Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
 OY 2 IALRYRILRY 12
 Db :|:| :|:|:
 24 LAVREKVLRY 34
 RESULT 5
 ABB01557
 ID ABB01557 standard; Peptide; 38 AA.
 AC ABB01557;
 XX 03-JAN-2002 (first entry)
 XX Viral core polypeptide, SEQ ID NO: 84.
 XX Human immunodeficiency virus; HIV; respiratory syncytial virus; RSV;
 XX virucide; heptad repeat region; transmembrane protein; gp41; HR1; HR2;
 XX infection.
 XX Viridiae.
 XX WO200164013-A2.
 XX PD 07-SEP-2001.
 XX XX 07-FEB-2001; 2001WO-US03988.
 XX PR 29-FEB-2000; 2000US-0515965.
 XX PA (TRIM-) TRIMERIS INC.
 XX PI Antczak JB, Delmedico MK, Erickson JB, Lambert DM, Sista P;
 XX DR WPI; 2001-514829/56.
 XX

PT Heptad repeat region peptide analogs useful for inhibiting virus/cells
 PT fusion, useful for treating HIV and Respiratory Syncytial Virus
 PT infection -
 XX XX Disclosure; Page 188; 587pp; English.
 XX CC The invention relates to isolated analogues of the heptad repeat region
 XX CC peptides DP178 and DP107. DP178 and DP107 correspond to amino acids
 XX CC 638-673 (heptad repeat region HR2) and 558-595 (heptad repeat region
 XX CC HR1) respectively, of HIV-1LAI transmembrane protein gp41. The HR1
 XX CC and HR2 regions of proteins interact non-covalently with each other
 XX CC and/or with peptides derived from them. This interaction is required for
 XX CC normal infectivity of viruses such as RSV and HIV. The heptad
 XX CC repeat region peptide analogues may be used to inhibit respiratory
 XX CC syncytial virus (RSV) infection in a cell. They may also be used to
 XX CC inhibit HIV infection. The present sequence is a peptide provided in
 XX CC the specification.
 XX SQ Sequence 38 AA;
 XX Query Match 56.9%; Score 33; DB 22; Length 38;
 XX Best Local Similarity 45.5%; Pred. No. 38;
 XX Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
 OY 2 IALRYRILRY 12
 Db :|:| :|:|:
 24 LAVREKVLRY 34
 RESULT 6
 AAU12641
 ID AAU12641 standard; Peptide; 38 AA.
 XX AC AAU12641;
 XX DT 21-NOV-2001 (first entry)
 XX DE DP178-like/DP107-like peptide T-88.
 XX KW Anti-retroviral; DP178-like; DP107-like; transmembrane protein gp41;
 XX KW antifuseogenic; antiviral; HIV transmission; mutant; mutain.
 XX OS Human immunodeficiency virus 1 isolate LAI.
 XX OS Synthetic.
 XX EH Key Location/Qualifiers
 XX FT Modified-site 1 /note= "N-terminal is substituted by Ac"
 XX FT Modified-site 38 /note= "C-terminal amide"
 XX FT WO200151673-A2.
 XX PN 19-JUL-2001.
 XX XX 05-JUL-2000; 2000WO-US35727.
 XX PR 09-JUL-1999; 99US-0350841.
 XX XX (TRIM-) TRIMERIS INC.
 XX XX Jeffs P, Lackey JW, Erickson JB, Lawless MK, Merutka G;
 XX DR WPI; 2001-442157/47.
 XX PT Identifying a compound that inhibits the formation of or disrupts a
 XX PT DP107/DP178 complex, especially compounds with antifuseogenic, antiviral
 XX PT or intracellular modulatory activity, by detecting the formation of a
 XX PT DP107/DP178 complex -
 XX XX Disclosure; Page 54; 259pp; English.
 XX CC The present invention relates to peptides which exhibit anti-retroviral

CC activity. The peptides of the invention (AAU12559-AAU14009) comprise
CC DP178-like and DP107-like peptides. The DP178 peptide corresponds
CC to amino acids 639-673 of the transmembrane protein gp41 from human
CC immunodeficiency virus 1 (HIV-1) isolate LA1. The DP107 peptide
CC corresponds to amino acids 558-595 of gp41 from HIV-1LA1. The invention
CC also relates to a method of identifying compounds that inhibit the
CC formation of or disrupt a DP107/DP178 complex. The method comprises
CC detecting the formation of a DP107/DP178 complex, both in the presence
CC or absence of a test compound, in a reaction mixture containing DP107
CC and DP178 peptides. The method is useful for identifying compounds,
CC including small molecule compounds, which may themselves exhibit
CC anti-fusogenic, antiviral or intracellular modulatory activity. The
CC DP178-like/DP107-like peptides are useful to inhibit human and non-human
CC retroviral, particularly HIV, transmission to uninfected cells. The
CC present sequence represents one of the DP178-like/DP107-like peptides
CC of the invention.
XX
SQ Sequence 38 AA;

Query Match 56.9%; Score 33; DB 22; Length 38;
Best Local Similarity 45.5%; Pred. No. 38;
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

OY 2 IALRYRILLRY 12
Db 24 LAVREKVLIRY 34

RESULT 7
AAB77088
ID AAB77088 standard; Peptide; 38 AA.
XX AAB77088;
XX
XX 19-APR-2001 (first entry)
DT
XX
DE Core polypeptide T88.
XX
XX Core polypeptide; enhancer; antiviral; anti-HIV;
KW vinnicide; hepatocytic; antiinflammatory; hybrid polypeptide;
KW coiled-coil peptide interaction; fusion-related disorder;
KW bacterial infection; viral infection.
XX
XX Unidentified.
OS
XX WO200103723-A1.
PN
XX 18-JAN-2001.
PD
XX 10-JUL-2000; 2000WO-US18772.
PF
XX 09-JUL-1999; 99US-0350641.
PR
XX (TRIM-) TRIMERIS INC.
PA
XX Barney S, Guthrie KI, Merutka G, Anwer MK, Lambert DM,
PI
XX WPI; 2001-147136/15.
DR
XX
PT New hybrid polypeptide, useful for preventing, treating and diagnosing
PT e.g. viral infections, comprises an enhancer peptide linked to a core
PT polypeptide -
PS
XX Disclosure; Page 31; 151dp; English.
XX
XX The present sequence is a core polypeptide which may be linked to
CC an enhancer peptide to form a novel hybrid polypeptide. The hybrid
CC polypeptide exhibits enhanced pharmacokinetic properties relative to
CC those exhibited by the core polypeptide when introduced into a living
CC system. It is used to increase the in vitro or ex vivo half-life of
CC the core polypeptide. The hybrid and core polypeptides can be used for
CC modulating fusogenic events and intracellular processes involving
CC coiled-coil peptide interactions. Other uses include preventing,

CC treating and/or diagnosing disorders involving fusion events (e.g.
CC modulation of neurotransmitter exchange and sperm-egg fusion),
CC intracellular and viral infections involving coiled-coil peptides (e.g. bacterial
CC infections) and viral infections that involve cell-cell and/or
CC virus-cell fusion (e.g. viral infections caused by human
CC immunodeficiency virus, respiratory syncytial virus, Epstein-Barr
CC virus, hepatitis B virus, Mason-Pfizer virus and polio virus).
CC The enhancer peptide sequence increases the half-life and reduces the
CC clearance rate of therapeutic peptides, which increases their efficacy
CC and minimises the incidence and severity of adverse side effects.
CC In addition, this increases the sensitivity of the diagnostic procedure
CC in which they are used.
XX
SQ Sequence 38 AA;

Query Match 56.9%; Score 33; DB 22; Length 38;
Best Local Similarity 45.5%; Pred. No. 38;
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

OY 2 IALRYRILLRY 12
Db 24 LAVREKVLIRY 34

RESULT 8
AAM74957
ID AAM74957 standard; Protein; 39 AA.
XX AAM74957;
XX
XX 19-JAN-1999 (first entry)
DT
XX
DE Human secreted protein encoded by gene 79 clone HSKNE46.
XX
XX Human; secreted protein; testis; tumour; foetal brain tissue;
KW fusion protein; cancer; central nervous system; seizure;
KW diagnosis; neurodegenerative disease.
XX
XX Homo sapiens.
OS
XX WO9839448-A2.
PN
XX 11-SEP-1998.
PD
XX 06-MAR-1998; 98WO-US04493.
PF
XX 02-OCT-1997; 97US-0061060.
PR 07-MAR-1997; 97US-0038621.
PR 07-MAR-1997; 97US-0040161.
PR 07-MAR-1997; 97US-0040162.
PR 07-MAR-1997; 97US-0040163.
PR 07-MAR-1997; 97US-0040333.
PR 07-MAR-1997; 97US-0040334.
PR 07-MAR-1997; 97US-0040336.
PR 07-MAR-1997; 97US-0040626.
PR 11-APR-1997; 97US-0043311.
PR 11-APR-1997; 97US-0043312.
PR 11-APR-1997; 97US-0043313.
PR 11-APR-1997; 97US-0043314.
PR 11-APR-1997; 97US-0043568.
PR 11-APR-1997; 97US-0043569.
PR 11-APR-1997; 97US-0043576.
PR 11-APR-1997; 97US-0043578.
PR 11-APR-1997; 97US-0043580.
PR 11-APR-1997; 97US-0043669.
PR 11-APR-1997; 97US-0043670.
PR 11-APR-1997; 97US-0043671.
PR 11-APR-1997; 97US-0043672.
PR 11-APR-1997; 97US-0043674.
PR 23-MAY-1997; 97US-0047492.
PR 23-MAY-1997; 97US-0047500.
PR 23-MAY-1997; 97US-0047501.
PR 23-MAY-1997; 97US-0047502.

PR 23-MAY-1997; 97US-0047503.
 PR 23-MAY-1997; 97US-0047581.
 PR 23-MAY-1997; 97US-0047582.
 PR 23-MAY-1997; 97US-0047583.
 PR 23-MAY-1997; 97US-0047584.
 PR 23-MAY-1997; 97US-0047585.
 PR 23-MAY-1997; 97US-0047586.
 PR 23-MAY-1997; 97US-0047587.
 PR 23-MAY-1997; 97US-0047588.
 PR 23-MAY-1997; 97US-0047589.
 PR 23-MAY-1997; 97US-0047590.
 PR 23-MAY-1997; 97US-0047592.
 PR 23-MAY-1997; 97US-0047593.
 PR 23-MAY-1997; 97US-0047594.
 PR 23-MAY-1997; 97US-0047595.
 PR 23-MAY-1997; 97US-0047596.
 PR 23-MAY-1997; 97US-0047597.
 PR 23-MAY-1997; 97US-0047598.
 PR 23-MAY-1997; 97US-0047599.
 PR 23-MAY-1997; 97US-0047600.
 PR 23-MAY-1997; 97US-0047601.
 PR 23-MAY-1997; 97US-0047612.
 PR 23-MAY-1997; 97US-0047613.
 PR 23-MAY-1997; 97US-0047614.
 PR 23-MAY-1997; 97US-0047615.
 PR 23-MAY-1997; 97US-0047617.
 PR 23-MAY-1997; 97US-0047618.
 PR 23-MAY-1997; 97US-0047632.
 PR 23-MAY-1997; 97US-0047633.
 PR 06-JUN-1997; 97US-0048964.
 PR 06-JUN-1997; 97US-0048974.
 PR 13-JUN-1997; 97US-0049610.
 PR 08-JUL-1997; 97US-0051926.
 PR 16-JUL-1997; 97US-0052874.
 PR 18-AUG-1997; 97US-0055724.
 PR 22-AUG-1997; 97US-0056630.
 PR 22-AUG-1997; 97US-0056631.
 PR 22-AUG-1997; 97US-0056632.
 PR 22-AUG-1997; 97US-0056633.
 PR 22-AUG-1997; 97US-0056637.
 PR 22-AUG-1997; 97US-0056662.
 PR 22-AUG-1997; 97US-0056664.
 PR 22-AUG-1997; 97US-0056845.
 PR 22-AUG-1997; 97US-0056862.
 PR 22-AUG-1997; 97US-0056864.
 PR 22-AUG-1997; 97US-0056872.
 PR 22-AUG-1997; 97US-0056874.
 PR 22-AUG-1997; 97US-0056875.
 PR 22-AUG-1997; 97US-0056876.
 PR 22-AUG-1997; 97US-0056877.
 PR 22-AUG-1997; 97US-0056878.
 PR 22-AUG-1997; 97US-0056879.
 PR 22-AUG-1997; 97US-0056880.
 PR 22-AUG-1997; 97US-0056881.
 PR 22-AUG-1997; 97US-0056882.
 PR 22-AUG-1997; 97US-0056884.
 PR 22-AUG-1997; 97US-0056886.
 PR 22-AUG-1997; 97US-0056887.
 PR 22-AUG-1997; 97US-0056888.
 PR 22-AUG-1997; 97US-0056889.
 PR 22-AUG-1997; 97US-0056892.
 PR 22-AUG-1997; 97US-0056893.
 PR 22-AUG-1997; 97US-0056894.
 PR 22-AUG-1997; 97US-0056903.
 PR 22-AUG-1997; 97US-0056908.
 PR 22-AUG-1997; 97US-0056909.
 PR 22-AUG-1997; 97US-0056910.
 PR 22-AUG-1997; 97US-0056911.
 PR 05-SEP-1997; 97US-0057650.
 PR 05-SEP-1997; 97US-0057659.
 PR 12-SEP-1997; 97US-0057761.
 PR 12-SEP-1997; 97US-0058785.
 PR XX

PA (HUMA-) HUMAN GENOME SCI INC.
 XX Bednarik DP, Brewer LA, Carter KC, Duan R, Ebner R, Endress GA;
 PI Feng P, Ferrie AM, Fischer CL, Florence KA, Greene JM, Hu JS;
 PI Kyaw H, Lafleur DW, Li Y, Moore PA, Ni J, Olsen HS, Rosen CA;
 PI Ruben SM, Shi Y, Soppet DR, Young PE, Yu GL, Zeng Z;
 XX WPI; 1998-506364/43.
 DR N-PSDB; AAV59742.
 XX
 PT New isolated human genes and the secreted polypeptide(s) they encode
 PT - useful for diagnosis and treatment of e.g. cancers, neurological
 PT disorders, immune diseases, inflammation or blood disorders
 XX
 PS Claim 1; Page 674; 721pp; English.
 XX
 CC This sequence represents a secreted human protein encoded by the nucleic
 CC acid molecule designated Gene 79 from the human cDNA clone HSKNE46
 CC (deposited as clone ATCC 97900 and ATCC 209046).
 CC The gene can be used to generate fusion proteins by linking to the gene
 CC to a human immunoglobulin Fc portion (e.g. AAV59502) for increasing the
 CC stability of the fused protein as compared to the human protein only.
 CC The invention relates to 186 novel genes and their fragments (nucleic
 CC acid sequences: AAV59511-V59812; amino acid sequences AAW74731-W75026)
 CC which are useful for preventing, treating or ameliorating medical
 CC conditions e.g. by protein or gene therapy. Also, pathological
 CC conditions can be diagnosed by determining the amount of the new
 CC polypeptides in a sample or by determining the presence of mutations in
 CC the new polynucleotides. Specific uses are described for each of the 186
 CC polynucleotides, based on which tissues they are most highly expressed in
 CC (see AAV59511 for described uses).
 XX
 SQ Sequence 39 AA;
 Query Match 55.2%; Score 32; DB 19; Length 39;
 Best Local Similarity 77.8%; Pred. No. 59;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 RIALRYRIL 9
 Db 23 RIALYRIV 31
 |||||
 |||||
 RESULT 9
 AAM61017
 ID AAM61017 standard; Protein; 54 AA.
 XX
 AC AAM61017;
 XX
 XX 05-NOV-2001 (first entry)
 XX Human brain expressed single exon probe encoded protein SEQ ID NO: 33122.
 DE Human; brain expressed exon; gene expression analysis; probe;
 KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
 KW epilepsy; cancer.
 XX
 OS Homo sapiens.
 XX
 FN WO200157275-A2.
 XX
 PD 09-AUG-2001.
 XX
 XX 30-JAN-2001; 2001WO-US00667.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 PR XX

```

PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-483446/52.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
XX brains -
XX
XX Example 4; SEQ ID NO: 33122; 650pp + Sequence Listing; English.
XX
XX
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX brain. They can be used to measure gene expression in brain cell samples,
XX which may enable the diagnosis and improved treatment of nervous system
XX diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
XX epilepsy and cancers. The present sequence is a protein encoded by one of
XX the probes of the invention.
XX
XX Sequence 54 AA:
XX
XX
XX Query Match 55.2%; Score 32; DB 22; Length 54;
XX Best Local Similarity 70.0%; Pred. No. 83;
XX Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0
XX
XX 1 RIALRYRILL 10
XX | : | : | : |
XX 24 RIALRILL 33
XX
XX RESULT 10
XX AAM73714
XX ID AAM73714 standard; Protein; 54 AA.
XX AC AAM73714;
XX DT 06-NOV-2001 (first entry)
XX
XX Human bone marrow expressed probe encoded protein SEQ ID NO: 34020.
XX
XX Human; bone marrow expressed exon; gene expression analysis; probe;
XX microarray; cancer; leukaemia; lymphoma; myeloma.
XX
XX Homo sapiens.
XX
XX MO200157276-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00668.
XX
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488900/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human bone marrow -
XX
XX Example 4; SEQ ID NO: 34020; 658bp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX bone marrow. They can be used to measure gene expression in bone marrow
XX

```

CC	samples, which may enable the improved diagnosis and treatment of cancers
CC	such as lymphoma, leukemia and myeloma. The present sequence is a
CC	protein encoded by one of the probes of the invention.
XX	
SQ	Sequence 54 AA;
QY	1 RIALRYRILL 10 : : :
DB	24 RIALHRRILL 33
RESULT 11	
AAM33900	
ID	AAM33900 standard; Protein; 54 AA.
XX	
AC	AAM33900;
XX	
DT	17-OCT-2001 (first entry)
XX	
DE	Peptide #937 encoded by probe for measuring placental gene expression.
XX	
KW	Probe; microarray; human; placenta; antenatal diagnosis;
KW	genetic disorder.
XX	
OS	Homo sapiens.
XX	
PN	WO200157272-A2.
XX	
PD	09-AUG-2001.
XX	
PF	30-JAN-2001; 2001WO-US00663.
XX	
PR	04-FEB-2000; 2000US-0180312.
PR	26-MAY-2000; 2000US-0207456.
PR	30-JUN-2000; 2000US-0608408.
PR	03-AUG-2000; 2000US-0632366.
PR	21-SEP-2000; 2000US-0234687.
PR	27-SEP-2000; 2000US-0236359.
PR	04-OCT-2000; 2000GB-0024263.
XX	
PA	(MOLE-) MOLECULAR DYNAMICS INC.
XX	
PI	Penn SG, Hanzel DK, Chen W, Rank DR;
XX	
DR	WPI; 2001-488897/53.
XX	
PT	Human genome-derived single exon nucleic acid probes useful for
PT	analyzing gene expression in human placenta -
XX	
PS	Claim 27; SEQ ID No 34169; 654pp; English.
XX	
XX	
CC	The present invention relates to single exon nucleic acid probes (SENPs;
CC	see AAI31315-AA157546). The present sequence is a peptide encoded by one
CC	such probe. The probes are useful for producing a microarray for
CC	predicting, measuring and displaying gene expression in samples derived
CC	from human placenta. The probes are useful for antenatal diagnosis of
CC	human genetic disorders.
XX	
SQ	Sequence 54 AA;
QY	Query Match 55.2%; Score 32; DB 22; Length 54;
DB	Best Local Similarity 70.0%; Pred. No. 83;
Matches	7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
Matches	7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY	1 RIALRYRILL 10 : : :
DB	24 RIALHRRILL 33

RESULT 12.

ABG43591
ID ABG43591 standard; Peptide; 54 AA.
AC ABG43591;
XX
DT 19-AUG-2002 (first entry)
XX
DE Human peptide encoded by genome-derived single exon probe SEQ ID 33256.
XX
KW Human; single exon probe; asthma; lung cancer; COPD; ILD;
KW chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemorrhoidosis;
KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesia; pulmonary hypertension;
KW hyaline membrane disease.
XX
OS Homo sapiens.
XX
PN WO200186003-A2.
XX
PD 15-NOV-2001.
XX
PF 30-JAN-2001; 2001WO-US00665.
XX
PR 04-FEB-2000; 2000US-180312P.
PR 26-MAY-2000; 2000US-207456P.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-234687P.
PR 27-SEP-2000; 2000US-236359P.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2002-114183/15.
XX
PT Spatially-addressable set of single exon nucleic acid probes, used to
PT measure gene expression in human lung samples -
XX
PS Claim 27; SEQ ID No 33256; 634pp; English.
XX
CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human lung comprising single exon nucleic acid probes having one of
CC 12614 nucleic acid sequences mentioned in the specification, or their
CC complements or the 12387 open reading frames derived from the 12614
CC probes. Also included are a microarray comprising the novel set of
CC probes; the novel set of probes which hybridise at high stringency to a
CC nucleic acid expressed in the human lung; measuring gene expression in a
CC sample derived from human lung, comprising (a) contacting the array with
CC a collection of detectably labeled nucleic acids derived from human lung
CC mRNA, and (b) measuring the label detectably bound to each probe of
CC the array; identifying exons in a eukaryotic genome, comprising
CC (a) algorithmically predicting at least one exon from genomic sequences
CC of the eukaryote; and (b) detecting specific hybridisation of detectably
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray; assigning exons to a single gene,
CC comprising (a) identifying exons from genomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridisation to a single exon
CC microarrays having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene
CC expression analysis, and for identifying exons in a gene, particularly

CC using human lung derived mRNA and for the study of lung diseases
CC such as asthma, lung cancer, chronic obstructive pulmonary disease
CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
CC haemorrhoidosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
CC pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic
CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
CC and hyaline membrane disease. The present sequence is a peptide/protein
CC encoded by a single exon probe of the invention.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 54 AA;

Query Match 55.2%; Score 32; DB 23; Length 54;
Best Local Similarity 70.0%; Pred. No. 83;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RIALRYRILL 10
Db 24 RLALLHRIILL 33

RESULT 13

AAR95429
ID AAR95429 standard; peptide; 12 AA.

AC AAR95429;

DT 12-NOV-1996 (first entry)

DE HLA-B2702 84-79-84 palindrome.

XX HLA; p74; alpha-helix; human-leucocyte-associated antigen; inhibitor;
KW T-cell lysate; membrane protein; mammal; heat shock protein; Hsc70; APC;
KW B cell; calcium influx; cytotoxic T lymphocyte; CTL; differentiation;
KW cytotoxicity; antigen presenting cell.
XX
OS Synthetic.

XX WO9513288-A1.
XX
XX 18-MAY-1995.

XX 10-NOV-1994; 94WO-US12985.

XX 10-NOV-1993; 93US-0150493.

XX (STRD) UNIV LELAND STANFORD JUNIOR.

XX Clayberger C, Krensky AM;

XX WPI; 1995-194027/25.

XX Compens. comprising lymphoid surface membrane proteins - which may
XX inhibit cytolytic activity and differentiation of CTLs.
XX Example; Page 12; 29pp; English.

XX AAR95413, and AAR95415-R95431 represent palindromes and fragments of
XX human-leucocyte-associated antigens. This sequence represents the
XX HLA-B2702 84-79-84 palindrome. These sequences can be used to isolate
XX the protein p74 from a T-cell lysate. p74 is a T-cell surface membrane
XX protein associated with T-cell activation in mammalian T-cells, and is
XX also immunologically cross reactive with the heat shock protein Hsc70.
XX p74 is found in a limited number of cell types, but is particularly
XX expressed on B and T cells. p74 can be isolated by lysis of a suitable
XX cell with an amphoteric detergent, and then passed through an affinity
XX column containing a covalently bound HLA-B2702 palindromic peptide.
XX Compositions comprising the extracellular fragment of p74 combined with

CC HLA-B*2702.60-.84 (see AAR95416), induces calcium influx, and inhibits
 CC cytotoxic T lymphocyte (CTL) differentiation or cytolysis. Candidate
 CC compounds can be screened for their effect on the cytolytic activity of
 CC T-cells, by combining them with the extracellular portion of p74 and
 CC determining the amount of binding between the candidate compound and
 CC p74. Modulation of CTL activity can be inhibited in a cellular
 CC composition containing T-cells and antigen presenting cells (APCs), by
 CC adding to the mix the extracellular portion of p74, in an amount
 CC sufficient to compete with p74 for the binding of the p74 ligand.
 CC XX

SQ Sequence 12 AA;

Query Match 54.3%; Score 31.5; DB 16; Length 12;
 Best Local Similarity 66.7%; Pred. No. 21;
 Matches 8; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

OY 1 RIALRYRLRY 12
 Db 2 RLAIR-RLAIRY 12

RESULT 14

AAW33798 ID AAW33798 standard; peptide; 12 AA.

AC AAW33798;

DT 19-JUN-1998 (first entry)

DE Peptide B2702.84-79/79-84 tested for immunomodulating activity.

XX Immunomodulating dimer; immunosuppressant drug; CTL activation;
 KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;
 KW rejection.

OS Synthetic.

OS Homo sapiens.

XX W09744351-A1.

PD 27-NOV-1997.

PE 22-MAY-1997; 97WO-US08689.

PR 24-MAY-1996; 96US-0653294.

PA (STRD) UNIV LELAND STANFORD JUNIOR.

PI Beulow R, Clayberger C, Krensky AM;

XX WPI; 1998-086530/08.

PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B
 alpha-1 domain, used for preventing rejection of transplants or
 PT treating autoimmune diseases

XX Example 1; Page 19; 41pp; English.

PS Peptides AAW33784-98 and AAW33778-9 were assayed for their
 CC immunomodulating activity. A peptide-type compound or variant is claimed
 CC which has immunomodulating activity, including the N-terminal acylated
 CC and/or C-terminal amidated or esterified forms of up to 60 amino acids,
 CC where the peptide-type compound comprises the formula: A-B, where A, B =
 CC (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = B or V; aa77 =
 CC D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a hydrophobic or
 CC small amino acid; aa82 = R or L; aa83 = G or R; and aa represents amino
 CC acid. The sequence in the brackets may optionally be absent or truncated
 CC at any peptide type bond within the brackets. The compounds comprise
 CC amino acid sequences related to a Class I HLA-B alpha1 domain (positions
 CC 79-84). They can be used to inhibit cytotoxic T-lymphocytes (CTL) from
 CC undesirably attacking cells in a host or in vitro. They can also be
 CC used in combination with antigenic peptides or proteins of interest to
 CC activate CTLs. They can also inhibit the proliferation of T cells in

CC response to anti-CD3. The peptide can be used for preventing rejection
 CC of transplants or for treating autoimmune diseases, e.g. diabetes,
 CC rheumatoid arthritis and lupus erythematosus. The products can also be
 CC used for detection and diagnosis.
 CC XX

SQ Sequence 12 AA;

Query Match 54.3%; Score 31.5; DB 19; Length 12;
 Best Local Similarity 66.7%; Pred. No. 21;
 Matches 8; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

OY 1 RIALRYRLRY 12
 Db 2 RLAIR-RLAIRY 12

RESULT 15

AAW33798 ID AAW33798 standard; peptide; 12 AA.

AC AAW33798;

DT 19-JUN-1998 (first entry)

DE Immunomodulating dimer peptide #3.

XX Immunomodulating dimer; immunosuppressant drug; CTL activation;
 KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;
 KW rejection.

OS Synthetic.

OS Homo sapiens.

XX W09744351-A1.

PD 27-NOV-1997.

PE 22-MAY-1997; 97WO-US08689.

PR 24-MAY-1996; 96US-0653294.

PA (STRD) UNIV LELAND STANFORD JUNIOR.

PI Beulow R, Clayberger C, Krensky AM;

XX WPI; 1998-086530/08.

PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B
 alpha-1 domain, used for preventing rejection of transplants or
 PT treating autoimmune diseases

XX Claim 17; Page 35; 41pp; English.

PS This sequence represents a specifically claimed immunomodulating
 CC dimer peptide of the invention. A peptide-type compound or variant is
 CC claimed which has immunomodulating activity, including the N-terminal
 CC acylated and/or C-terminal amidated or esterified forms of up to 60
 CC amino acids, where the peptide-type compound comprises the formula: A-B,
 CC where A, B = (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = B or
 CC V; aa77 = D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a
 CC hydrophobic or small amino acid; aa82 = R or L; aa83 = G or R; and aa
 CC represents amino acid. The sequence in the brackets may optionally be
 CC absent or truncated at any peptide type bond within the brackets. The
 CC compounds comprise amino acid sequences related to a Class I HLA-B
 CC alpha1 domain (positions 79-84). They can be used to inhibit cytotoxic
 CC T-lymphocytes (CTL) from undesirably attacking cells in a host or in
 CC vitro. They can also be used in combination with antigenic peptides or
 CC proteins of interest to activate CTLs. They can also inhibit the
 CC proliferation of T cells in response to anti-CD3. The peptide can be
 CC used for preventing rejection of transplants or for treating autoimmune
 CC diseases, e.g. diabetes, rheumatoid arthritis and lupus erythematosus.
 CC The products can also be used for detection and diagnosis.

SQ Sequence 12 AA;

Query Match 54.3%; Score 31.5; DB 19; Length 12;
 Best Local Similarity 66.7%; Pred. No. 21;
 Matches 8; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

Qy 1 RIALRYRILLRY 12

Db 2 RLAIR-RIALRY 12

Search completed: December 19, 2002, 16:19:08
 Job time : 29.5882 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 19, 2002, 16:14:32 ; Search time 21.1765 Seconds
(without alignments)
116.760 Million cell updates/sec

Title: US-08-653-294C-41

Perfect score: 58

Sequence: 1 RIALRYILLRY 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 45785

Minimum DB seq length: 0

Maximum DB seq length: 60

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL.21.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaea.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30	51.7	42	11 O54741	O54741 mus musculus
2	29	50.0	18	11 O9QVSA	O9QVSA rattus sp.
3	29	50.0	54	16 Q9PBA2	Q9PBA2 xyella fas
4	28	48.3	53	16 Q9KCES	Q9KCES bacillus ha
5	28	48.3	54	12 Q91G03	Q91G03 chilo iride
6	28	48.3	55	10 Q94FX9	Q94FX9 zea mays (m
7	28	48.3	55	10 Q93VN4	Q93VN4 zea mays (m
8	28	48.3	58	16 Q92HR4	Q92HR4 rickettsia
9	28	48.3	60	3 Q9HFH8	Q9HFH8 pichia etch
10	28	48.3	60	10 P93387	P93387 nicotiana t
11	27	46.6	39	12 Q9PXR1	Q9PXR1 tobacco mos
12	27	46.6	47	7 Q98155	Q98155 mirounga le
13	27	46.6	47	7 Q98156	Q98156 mirounga le
14	27	46.6	47	7 Q98157	Q98157 mirounga le
15	27	46.6	47	7 Q98159	Q98159 mirounga le
16	27	46.6	50	9 O34049	O34049 streptococc

17	27	46.6	51	2 Q9EZE2	Q9EZE2 escherichia
18	27	46.6	51	9 Q9XJU7	Q9XJU7 streptococc
19	27	46.6	52	16 Q97M11	Q97M11 clostridium
20	26.5	45.7	59	7 Q80094	Q80094 homo sapien
21	26	44.8	32	3 Q8TGT3	Q8TGT3 saccharomyc
22	26	44.8	42	8 Q9MSG6	Q9MSG6 plasmodium
23	26	44.8	47	7 Q98158	Q98158 mirounga le
24	26	44.8	47	7 Q98160	Q98160 mirounga le
25	26	44.8	47	7 Q98161	Q98161 mirounga le
26	26	44.8	47	7 Q98168	Q98168 arctocephal
27	26	44.8	50	15 Q990A8	Q990A8 human immun
28	26	44.8	54	3 Q07058	Q07058 saccharomyc
29	25	43.1	25	6 Q9GK82	Q9GK82 ceratotheri
30	25	43.1	29	16 Q9JXD6	Q9JXD6 neisseria m
31	25	43.1	44	12 Q91G14	Q91G14 chilo iride
32	25	43.1	45	12 Q91FUI	Q91FUI chilo iride
33	25	43.1	47	15 Q9QLP6	Q9QLP6 human immun
34	25	43.1	48	4 Q9UML1	Q9UML1 homo sapien
35	25	43.1	50	16 Q9XIC2	Q9XIC2 thermotoga
36	25	43.1	50	17 Q97225	Q97225 sulfolobus
37	25	43.1	51	2 Q57438	Q57438 lactobacill
38	25	43.1	51	5 Q25800	Q25800 plasmodium
39	25	43.1	51	9 Q9MCL8	Q9MCL8 streptococc
40	25	43.1	53	2 Q9AQD4	Q9AQD4 lactococcus
41	25	43.1	55	10 Q93WH7	Q93WH7 zea mays (m
42	25	43.1	56	10 Q94FX8	Q94FX8 zea mays (m
43	25	43.1	57	16 Q9PGF3	Q9PGF3 xyella fas
44	25	43.1	58	16 Q97HQ3	Q97HQ3 clostridium
45	25	43.1	58	16 Q9CGN4	Q9CGN4 lactococcus

ALIGNMENTS

RESULT 1
O54741 PRELIMINARY; PRT; 42 AA.

AC O54741;
DC 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Laminin 5 alpha3C chain (Fragment).
GN LAMA3 OR LAMA3.

OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=SV129;

RX MEDLINE=97400527; PubMed=9252362;

RA Ferrigno O., Vitrolle T., Galliano M.P., Chauvin N., Ortonne J.P.,

RT Meneguzzi G., Aberdam D.;

RL "Murine laminin 5 a3a and a3B isoform chains are generated by usage of

two promoters and alternative splicing."

RL J. Biol. Chem. 272:20502-20508(1997).

DR EMBL; Y08850; CAA70073.1; -

DR MBL; GGI:99909; Lama3.

FT NON TER 42 42

SQ SEQUENCE 42 AA; 4872 MW; 64C9D8DE70807EAO CRC64;

Query Match 51.7%; Score 30; DB 11; Length 42;

Best Local Similarity 38.9%; Pred. No. 1.3e+02;

Matches 7; Conservative 5; Mismatches 0; Indels 6; Gaps 1;

QY 1 RIALR-----YRILLRY 12

|||||

DB 14 RLSVRQSSLSLFLVILRY 31

RESULT 2

Q9QVSA

ID Q9QVSA PRELIMINARY; PRT; 18 AA.

```

AC Q90VS4;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE Guanylate cyclase-A receptor (Fragment).
OS Rattus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC NCBI_TaxID=10118;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95237966; PubMed=7721418;
RA Miao Z.H., Song D.L., Douglas J.G., Chang C.H.;
RT "Mutational inactivation of the catalytic domain of guanylate cyclase-
RT A receptor.";
RL Hypertension 25:694-698(1995).
SQ SEQUENCE 18 AA; 2137 MW; 61026B2DA726629B CRC64;

Query Match 50.0%; Score 29; DB 11; Length 18;
Best Local Similarity 85.7%; Pred. No. 84;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Cy 3 ALRYRL 9
Db 7 ALRYRL 13

RESULT 3
Q9PBA2 PRELIMINARY; PRT; 54 AA.
ID Q9PBA2;
AC Q9PBA2;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Hypothetical protein Xf2242.
GN Xf2242.
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OC Xylella.
OC NCBI_TaxID=2371;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=9A5C;
RX MEDLINE=20365717; PubMed=10910347;
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Brites M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carer H.,
RA Coutinho L.B., Colombo C., Costa F.P., Costa M.C.R., Costa-Neto C.M.,
RA Facincini A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
RA Gantier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Graber A.,
RA Ho P.L., Hohnsels J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Krieger J.E., Kuramae E.B., Laigret F., Lambais M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Martino C.L.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorcello C.B.,
RA Moon D.H., Nagai M.A., Nascimento L.R.T.O., Netto L.E.S.,
RA Nhami A.Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Patris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.A.Jr., Pesquero J.B.,
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silveira J.F., Silvestri M.L.Z., Siqueira M.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Tsal S.M., Tsubako M.H.,
RA Valla da H., Van Sluys M.A., Verjowski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Melandri J., Setubal J.C.;
RT "The genome sequence of the plant pathogen Xylella fastidiosa.";
RL Nature 406:151-159(2000).
EMBL, AF004037; AAF85041.1; -.

```

```

KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 54 AA; 5826 MW; 2D3582F1E200B2AD CRC64;

Query Match 50.0%; Score 29; DB 16; Length 54;
Best Local Similarity 45.5%; Pred. No. 2.5e+02;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Cy 2 IALRYRL 12
Db 29 LTLNHAIRLY 39

RESULT 4
Q9KCE5 PRELIMINARY; PRT; 53 AA.
ID Q9KCE5;
AC Q9KCE5;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Hypothetical protein BH1626.
GN BH1626.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Bacillaceae; Bacillus.
OC NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Maui N.,
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331(2000).
DR EMBL; AP001512; BAB05345.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 53 AA; 6234 MW; 343F2BA731F5E63C CRC64;

Query Match 48.3%; Score 28; DB 16; Length 53;
Best Local Similarity 57.1%; Pred. No. 3.8e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Cy 5 RYRILR 11
Db 5 RYRVIR 11

RESULT 5
Q91G03 PRELIMINARY; PRT; 54 AA.
ID Q91G03;
AC Q91G03;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE 144R.
OS Chilo iridescent virus (CIV) (Insect iridescent virus type 6).
OC Viruses; dsDNA viruses, no RNA stage; Iridoviridae; Iridovirus.
OC NCBI_TaxID=10488;
RN [1]
RP SEQUENCE FROM N.A.
RX Delius H., Darai G., Flugel R.M.;
RT "DNA analysis of insect iridescent virus 6: evidence for circular
RT permutation and terminal redundancy.";
RL J. Virol. 49:609-614(1984).
RP SEQUENCE FROM N.A.
RX MEDLINE=86174607; PubMed=3959991;
RA Lotbacher de Ruiz H., Gelderblom H., Hofmann W., Darai G.;
RT "Insect iridescent virus type 6 induced toxic degenerative hepatitis
RT in mice.";
RL Med. Microbiol. Immunol. 175:43-53(1986).
EMBL, AF004037; AAF85041.1; -.

```

RP SEQUENCE FROM N.A.
RX MEDLINE=87321126; PubMed=2820141;
RA Schnitzler P., Soltan J.B., Fischer M., Reisner H., Scholz J.,
RA Delius H., Darai G.;
RT "Molecular cloning and physical mapping of the genome of insect
RT iridescent virus type 6: further evidence for circular permutation of
RL the viral genome.";
RL Virology 160:66-74 (1987).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=89073752; PubMed=3201750;
RA Fischer M., Schnitzler P., Delius H., Darai G.;
RT "Identification and characterization of the repetitive DNA element in
RT the genome of insect iridescent virus type 6.";
RL Virology 167:485-496 (1988).
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=92196996; PubMed=1549908;
RA Handermann M., Schnitzler P., Rosen-Wolff A., Raab K., Sonntag K.C.,
RA Darai G.;
RT "Identification and mapping of origins of DNA replication within the
RT DNA sequences of the genome of insect iridescent virus type 6.";
RL Virus Genes 6:19-32 (1992).
RN [6]
RP SEQUENCE FROM N.A.
RX MEDLINE=93118242; PubMed=1475907;
RA Sonntag K.C., Darai G.;
RT "Characterization of the third origin of DNA replication of the genome
RT of insect iridescent virus type 6.";
RL Virus Genes 6:333-342 (1992).
RN [7]
RP SEQUENCE FROM N.A.
RX MEDLINE=93260401; PubMed=8492091;
RA Stohwasser R., Raab K., Schnitzler P., Janssen W., Darai G.;
RT "Identification of the gene encoding the major capsid protein of
RT insect iridescent virus type 6 by polymerase chain reaction.";
RL J. Gen. Virol. 74:873-879 (1993).
RN [8]
RP SEQUENCE FROM N.A.
RX MEDLINE=94167241; PubMed=8121799;
RA Schnitzler P., Hug M., Handermann M., Janssen W., Koonin E.V.,
RA Delius H., Darai G.;
RT "Identification of genes encoding zinc finger proteins, non-histone
RT chromosomal HMG protein homologue, and a putative GTP phosphohydrolase
RT in the genome of Chilo iridescent virus.";
RL Nucleic Acids Res. 22:158-166 (1994).
RN [9]
RP SEQUENCE FROM N.A.
RX MEDLINE=94353641; PubMed=8073636;
RA Sonntag K.C., Schnitzler P., Koonin E.V., Darai G.;
RT "Chilo iridescent virus encodes a putative helicase belonging to a
RT distinct family within the 'DEAD/H' superfamily: implications for the
RT evolution of large DNA viruses.";
RL Virus Genes 8:151-158 (1994).
RN [10]
RP SEQUENCE FROM N.A.
RX MEDLINE=95213160; PubMed=7698884;
RA Sonntag K.C., Schnitzler P., Janssen W., Darai G.;
RT "Identification of the primary structure and the coding capacity of
RT the genome of insect iridescent virus type 6 between the genome
RT coordinates 0.310 and 0.347 (7990 bp).";
RL Intervirology 37:287-297 (1994).
RN [11]
RP SEQUENCE FROM N.A.
RX MEDLINE=94292906; PubMed=8021587;
RA Schnitzler P., Sonntag K.C., Muller M., Janssen W., Bugert J.J.,
RA Koonin E.V., Darai G.;
RT "Insect iridescent virus type 6 encodes a polypeptide related to the
RT largest subunit of eukaryotic RNA polymerase II.";
RL J. Gen. Virol. 75:1557-1567 (1994).
RN [12]
RP SEQUENCE FROM N.A.
RX MEDLINE=98141693; PubMed=9482589;

RA Bahr U., Tidona C.A., Darai G.;
RT "The DNA sequence of Chilo iridescent virus between the genome
RT coordinates 0.101 and 0.391; similarities in coding strategy between
RT insect and vertebrate iridoviruses.";
RL Virus Genes 15:235-245 (1997).
RN [13]
RP SEQUENCE FROM N.A.
RX MEDLINE=99125223; PubMed=9926400;
RA Muller K., Tidona C.A., Bahr U., Darai G.;
RT "Identification of a thymidylate synthase gene within the genome of
RT Chilo iridescent virus.";
RL Virus Genes 17:243-258 (1998).
RN [14]
RP SEQUENCE FROM N.A.
RX MEDLINE=99383793; PubMed=10456793;
RA Muller K., Tidona C.A., Darai G.;
RT "Identification of a gene cluster within the genome of Chilo
RT iridescent virus encoding enzymes involved in viral DNA replication
RT and processing.";
RL Virus Genes 18:243-264 (1999).
RN [15]
RP SEQUENCE FROM N.A.
RX MEDLINE=21342589; PubMed=11448171;
RA Jakob N.J., Muller K., Bahr U., Darai G.;
RT "Analysis of the First Complete DNA Sequence of an Invertebrate
RT Iridovirus: Coding Strategy of the Genome of Chilo Iridescent Virus.";
RL Virology 286:182-196 (2001).
RN [16]
RP SEQUENCE FROM N.A.
RA Jakob N.J., Muller K., Bahr U., Darai G.;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF303741; AAK82029.1; -
SQ SEQUENCE 54 AA; 6581 MW; 5E8172628A363A55 CRC64;
Query Match 48.3%; Score 28; DB 12; Length 54;
Best Local Similarity 41.7%; Pred. No. 3.9e+02;
Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
QY 1 RIALRYRILLRY 12
Db 31 RLKLFHLLLYKY 42
RESULT 6
Q94FX9 PRELIMINARY; PRT; 55 AA.
AC Q94FX9;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE T cytoplasm male sterility restorer factor 2 (fragment).
GN RF2.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoidae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. INBRED LINE Q66;
RX MEDLINE=21238656; PubMed=11340182;
RA Liu F., Cui X., Horner H.T., Weiner H., Schnable P.S.;
RT "Mitochondrial aldehyde dehydrogenase activity is required for male
RT fertility in maize.";
RL Plant Cell 13:1083-1078 (2001).
DR EMBL; AF318133; AAK60511.1; -
DR InterPro; IPR002086; Aldehyde_dehydr.
DR Pfam; PF00171; aldedh; 1.
FT NON_TER 1
FT NON_TER 55
SQ SEQUENCE 55 AA; 6464 MW; 99988CC37A025580 CRC64;
Query Match 48.3%; Score 28; DB 10; Length 55;

Best Local Similarity 75.0%; Pred. No. 4e+02; Indels 0; Gaps 0;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 5 RYRILRY 12
| | | | |
| | | | |
DB 2 RSRILRF 9

RESULT 7

O93VN4 PRELIMINARY; PRT; 55 AA.

AC O93VN4;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE T cytoplasm male sterility restorer factor 2 (Fragment).
GN RF2.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21238656; PubMed=11340182;
RA Liu F., Cui X., Horner H.T., Weiner H., Schnable P.S.;
RT "Mitochondrial aldehyde dehydrogenase activity is required for male
fertility in maize."
RL Plant Cell 13:1063-1078(2001).
DR EMBL; AF318130; AAK60508.1; -.
DR InterPro; IPR002086; Aldehyde_dehydr.
DR Pfam; PF00171; aldedh; 1.
FT NON TER 1 55
SQ SEQUENCE 55 AA; 6465 MW; 999682C37A025F20 CRC64;

Query Match 48.3%; Score 28; DB 10; Length 55;
Best Local Similarity 75.0%; Pred. No. 4e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 5 RYRILRY 12
| | | | |
| | | | |
DB 2 RSRILRF 9

RESULT 8

O92HR4 PRELIMINARY; PRT; 58 AA.

AC O92HR4;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Hypothetical protein RC0707.
GN RC0707.
OS Rickettsia conorii.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Rickettsiae; Rickettsia.
OX NCBI_TaxID=781;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MALISH 7;
RX MEDLINE=21442074; PubMed=11557893;
RA Ogata H., Audic S., Renesto-Audiffren P., Fournier P.-E., Barbe V.,
RA Samson D., Roux V., Cossart P., Weissbach J., Claverie J.-M.,
RA Raoult D.;
RT "Mechanisms of evolution in Rickettsia conorii and R. prowazekii";
RL Science 293:2093-2098(2001).
DR EMBL; AB008629; AAL03245.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 58 AA; 6883 MW; E1217AAC34C32D18 CRC64;

Query Match 48.3%; Score 28; DB 16; Length 58;

Best Local Similarity 58.3%; Pred. No. 4.2e+02; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 RIALRYRILRY 12
| | | | |
| | | | |
DB 43 RRAWRMFLLY 54

RESULT 9

O9HFB8 PRELIMINARY; PRT; 60 AA.

AC O9HFB8;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Hypothetical 7.1 kDa protein.
OS Pichia etchellsii.
OG Plasmid pPE1B.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Pichia.
OX NCBI_TaxID=136928;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CBS2011;
RX MEDLINE=21340829; PubMed=11447601;
RA Klansen R., Tonsidou L., Larsen M., Weinhardt F.;
RT "Genome organization of the linear cytoplasmic element pPE1B from
RT Pichia etchellsii."
RL Yeast 18:953-961(2001).
DR EMBL; AJ278986; CAC08223.1; -.
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 60 AA; 7129 MW; FB4320B9C8044D2B CRC64;

Query Match 48.3%; Score 28; DB 3; Length 60;
Best Local Similarity 50.0%; Pred. No. 4.3e+02;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 RIALRYRIL 10
| | | | |
| | | | |
DB 38 RYRILRYL 47

RESULT 10

P93387 PRELIMINARY; PRT; 60 AA.

AC P93387;
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE ORF.
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; easterids I; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98187248; PubMed=9526500;
RA Karer E.E., Beachy R.N., Holt C.A.;
RT "Cloning of tobacco genes that elicit the hypersensitive response."
RL Plant Mol. Biol. 36:681-690(1998).
DR EMBL; U66272; AAC49978.1; -.
SQ SEQUENCE 60 AA; 6800 MW; 3091FAA0484CEA2C CRC64;

Query Match 48.3%; Score 28; DB 10; Length 60;
Best Local Similarity 44.4%; Pred. No. 4.3e+02;
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 4 LRYRILRY 12
| | | | |
| | | | |
DB 18 RYRILRY 26

RESULT 11
ID Q9PXR1 PRELIMINARY; PRT; 39 AA.
AC Q9PXR1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE CHARGED ORF-X protein.
OS Tobacco mosaic virus.
OC Viruses; SERNA positive-strand viruses, no DNA stage; Tobamovirus.
OX NCBI_TaxID=12242;
RN [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=94114596; PubMed=8286438;
RA Morozov S.Yu., Denisenko O.N., Zelenina D.A., Fedorkin O.N.,
RA Solov'yev A.G., Maiss E., Casper R., Atabekov J.G.;
RT "A novel open reading frame in tobacco mosaic virus genome coding for
RT a putative small, positively charged protein.";
RL Biochimie 75:659-665(1993).
SQ SEQUENCE 39 AA; 4714 MW; 21686D1DB5B859F6 CRC64;

Query Match 46.6%; Score 27; DB 12; Length 39;
Best Local Similarity 36.4%; Pred. No. 4.3e+02;
Matches 4; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 2 IALRYRLLRY 12
| : : : : :
Db 9 IIRFKYVLY 19

RESULT 12
ID Q98155 PRELIMINARY; PRT; 47 AA.
AC Q98155;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE MHC class II antigen (fragment).
GN MHC-DQB.
OS Mirounga leonina (Southern elephant seal).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Pinnipedia; Phocidae; Mirounga.
OX NCBI_TaxID=9715;
RN [1]
RN SEQUENCE FROM N.A.
RA Hoelzel A.R., Stephens J.C., O'Brien S.J.;
RT "Molecular genetic diversity and evolution at the MHC DQB locus in
RT four species of pinnipeds.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF111032; AAC96347.1; -.
DR InterPro; IPR000353; MHC II beta.
DR Pfam; PF00969; MHC II beta; 1.
DR ProDom; PD000328; MHC II beta; 1.
KW Glycoprotein; MHC II; Transmembrane.
FT NON_TER 1 1
FT NON_TER 47 47
SQ SEQUENCE 47 AA; 5727 MW; 0E672D7A5F436208 CRC64;

Query Match 46.6%; Score 27; DB 7; Length 47;
Best Local Similarity 62.5%; Pred. No. 5.2e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 5 RYRILLRY 12
| : : : : :
Db 10 RVRVLTLY 17

RESULT 13
ID Q98156 PRELIMINARY; PRT; 47 AA.
AC Q98156;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)

DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE MHC class II antigen (fragment).
GN MHC-DQB.
OS Mirounga leonina (Southern elephant seal).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Pinnipedia; Phocidae; Mirounga.
OX NCBI_TaxID=9715;
RN [1]
RN SEQUENCE FROM N.A.
RA Hoelzel A.R., Stephens J.C., O'Brien S.J.;
RT "Molecular genetic diversity and evolution at the MHC DQB locus in
RT four species of pinnipeds.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF111033; AAC96348.1; -.
DR InterPro; IPR000353; MHC II beta.
DR Pfam; PF00969; MHC II beta; 1.
DR ProDom; PD000328; MHC II beta; 1.
KW Glycoprotein; MHC II; Transmembrane.
FT NON_TER 1 1
FT NON_TER 47 47
SQ SEQUENCE 47 AA; 5731 MW; 0E672D6FBB122308 CRC64;

Query Match 46.6%; Score 27; DB 7; Length 47;
Best Local Similarity 62.5%; Pred. No. 5.2e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 5 RYRILLRY 12
| : : : : :
Db 10 RVRVLTLY 17

RESULT 14
ID Q98157 PRELIMINARY; PRT; 47 AA.
AC Q98157;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE MHC class II antigen (fragment).
GN MHC-DQB.
OS Mirounga leonina (Southern elephant seal).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Pinnipedia; Phocidae; Mirounga.
OX NCBI_TaxID=9715;
RN [1]
RN SEQUENCE FROM N.A.
RA Hoelzel A.R., Stephens J.C., O'Brien S.J.;
RT "Molecular genetic diversity and evolution at the MHC DQB locus in
RT four species of pinnipeds.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF111034; AAC96349.1; -.
DR InterPro; IPR000353; MHC II beta.
DR Pfam; PF00969; MHC II beta; 1.
DR ProDom; PD000328; MHC II beta; 1.
KW Glycoprotein; MHC II; Transmembrane.
FT NON_TER 1 1
FT NON_TER 47 47
SQ SEQUENCE 47 AA; 5775 MW; 0E672D6BAF436208 CRC64;

Query Match 46.6%; Score 27; DB 7; Length 47;
Best Local Similarity 62.5%; Pred. No. 5.2e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 5 RYRILLRY 12
| : : : : :
Db 10 RVRVLTLY 17

RESULT 15
ID Q98159 PRELIMINARY; PRT; 47 AA.
AC Q98159;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)

```

DT 01-MAY-1999 (TrEMBLrel. 10, last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, last annotation update)
DE MHC class II antigen (Fragment).
GN MHC-DQB.
OS Mirounga leonina (Southern elephant seal).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Pinnipedia; Phocidae; Mirounga.
OX NCBI_TaxID=9715;
RN [1]
RP SEQUENCE FROM N.A.
RA Hoelzel A.R., Stephens J.C., O'Brien S.J.;
RT "Molecular genetic diversity and evolution at the MHC DQB locus in
RT four species of pinnipeds."
RU Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL, AF11036; AAC96351.1; -.
DR InterPro; IPR000353; MHC_II_beta.
DR Pfam; PF00969; MHC_II_beta; 1.
DR ProDom; PD000328; MHC_II_beta; 1.
KW Glycoprotein; MHC II; Transmembrane.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 47 AA; 5727 MW; 0E672D7A5F436208 CRC64;

Query Match 46.6%; Score 27; DB 7; Length 47;
Best Local Similarity 62.5%; Pred. No. 5, 2e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 5 RYRILRY 12
| : | |
Db 10 RYRVLTRY 17

```

Search completed: December 19, 2002, 16:21:55
 Job time : 23.1765 secs

GenCore version 5.1.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 19, 2002, 16:03:12 ; Search time 5.47059 Seconds
(without alignments)
90.980 Million cell updates/sec

Title: US-08-653-294C-41
Perfect score: 58
Sequence: 1 RIALRYRILRY 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 5116

Minimum DB seq length: 0
Maximum DB seq length: 60

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	29	50.0	44	1 YIM8_BPPH1	P10431 bacterioph
2	25	43.1	46	1 YLCG_BP82	Q37874 bacterioph
3	25	43.1	46	1 YLCG_SCOLI	Q47272 escherichia
4	25	43.1	49	1 Y195_BPT7	P03804 bacterioph
5	25	43.1	51	1 YCUA_LACCU	P35923 lactobacill
6	25	43.1	57	1 VSH_MUMMP2	P28084 mumps virus
7	25	43.1	57	1 VSH_MUMPA	P28085 mumps virus
8	25	43.1	57	1 VSH_MUMPB	P28082 mumps virus
9	25	43.1	57	1 VSH_MUMPR	P28087 mumps virus
10	24	41.4	26	1 Y560_HELPY	O25285 helicobacte
11	24	41.4	38	1 VPU_HV1A2	P05949 human immun
12	24	41.4	49	1 Y195_BPT3	P10306 bacterioph
13	24	41.4	59	1 RL32_BORBU	O51646 borrelia bu
14	23	39.7	29	1 YCX4_OOSI	P49830 odontella s
15	23	39.7	36	1 LHG_RHOVI	P47420 mycoplasma
16	23	39.7	37	1 RL36_MYCGE	P08803 human immun
17	23	39.7	43	1 VPU_HV1C4	P41537 musca domes
18	23	39.7	44	1 DIUH_MUSDO	P82373 diptoptera
19	23	39.7	46	1 DIUH_DIPPU	P23465 locusta mig
20	23	39.7	46	1 DIUH_LOCPM	P41538 periplaneta
21	23	39.7	46	1 DIUH_PERAM	P82707 zootermopsi
22	23	39.7	46	1 DIUH_ZOONE	P11924 aplysia cal
23	23	39.7	54	1 ELHB_APLCA	P11925 aplysia cal
24	23	39.7	54	1 ELHC_APLCA	P32035 cuscutea ref
25	23	39.7	55	1 YCX1_CUSRE	P29605 burkholderi
26	23	39.7	56	1 LIPT_BURCE	P28086 mumps virus
27	23	39.7	57	1 VSH_MUMPK	P31090 anabaena va
28	23	39.7	59	1 PSAE_ANAVA	P13571 oryctolagus
29	22	37.9	24	1 HPTA_RABIT	P13896 western equ
30	22	37.9	37	1 POLN_WEEV	O78433 guillardi
31	22	37.9	37	1 PSBY_GUTHI	P15794 bacterioph
32	22	37.9	39	1 COA2_BPPM2	P56619 tenebrio mo
33	22	37.9	47	1 DIU2_TENMO	

RESULT 1

YIM8_BPPH1 STANDARD; PRT; 44 AA.
AC P10431;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 01-MAR-1989 (Rel. 10, Last annotation update)
DE Hypothetical immunity region protein 8.
OS Bacteriophage phi-105.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
OC Lambda-like viruses.
OX NCBI_TaxID=10717;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86056972; PubMed=3934047;
RA Cully D.F., Garro A.J.;
RT "Nucleotide sequence of the immunity region of Bacillus subtilis bacteriophage phi 105: identification of the repressor gene and its mRNA and protein products.";
RL Gene 38:153-164(1985).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----
DR EMBL; M11920; ; NOT_ANNOTATED_CDS.
DR PIR; G24521; IMBP8.
KW Hypothetical protein.

SQ SEQUENCE 44 AA; 5574 MW; C440D1B134841832 CRC64;

Query Match 50.0%; Score 29; DB 1; Length 44;

Best Local Similarity 66.7%; Pred. No. 21;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RIALRYRIL 9

Db 15 RLALKVRIL 23

RESULT 2

YLCG_BP82 STANDARD; PRT; 46 AA.

AC Q37874;

DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Hypothetical 5.6 kDa protein in rusa 3'region (ORF45).

OS Bacteriophage 82.

OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
OC Lambda-like viruses.

OX NCBI_TaxID=10705;

RN [1]

```

RP SEQUENCE FROM N.A.
RX MEDLINE=96196428; PubMed=8648624;
RA Mandi A.A., Sharples G.J., Mandal T.N., Lloyd R.G.;
RT "Holliday junction resolvases encoded by homologous rusa genes in
RL Escherichia coli K-12 and phage 82.";
CC J. Mol. Biol. 257:561-573(1996).
CC -1- SIMILARITY: STRONG, TO E. COLI YLIG.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, X92587; CAA63331.1; -
CC DR Hypothetical protein.
CC SQ SEQUENCE 46 AA; 5556 MW; E96413D0D5C6A265 CRC64;
-----
Query Match 43.1%; Score 25; DB 1; Length 46;
Best Local Similarity 71.4%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
-----
QY 6 YRILRY 12
DB 31 YRILKNY 37
-----
RESULT 3
YLIG_ECOLI STANDARD; PRT; 46 AA.
AC Q47272;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein ylig in lambdaoid Dlp12 prophage region.
GN YLIG OR B0550.1.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OC NCBI_TaxID=562;
OX [1]
RP SEQUENCE FROM N.A.
RX STRAIN=K12 / MG1655;
RX MEDLINE=96196428; PubMed=8648624;
RA Mandi A.A., Sharples G.J., Mandal T.N., Lloyd R.G.;
RT "Holliday junction resolvases encoded by homologous rusa genes in
RL Escherichia coli K-12 and phage 82.";
CC J. Mol. Biol. 257:561-573(1996).
CC (2)
CC SEQUENCE FROM N.A.
CC STRAIN=K12 / MG1655;
CC MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.C., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
CC Science 277:1453-1474(1997).
CC -1- SIMILARITY: STRONG, TO EQUIVALENT PROTEIN IN PHAGE 82.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, X92587; CAA63332.1; -
CC DR EMBL, AB000160; -; NOT_ANNOTATED_CDS.
CC Ecogene, BG14354; YLIG.

```

```

KM Hypothetical protein; Complete proteome.
SQ SEQUENCE 46 AA; 5493 MW; E5D339EBC756A26E CRC64;
-----
Query Match 43.1%; Score 25; DB 1; Length 46;
Best Local Similarity 71.4%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
-----
QY 6 YRILRY 12
DB 31 YRILKNY 37
-----
RESULT 4
YL95_BPT7 STANDARD; PRT; 49 AA.
AC P03804;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-MAR-1989 (Rel. 10, Last annotation update)
DE Hypothetical gene 19.5 protein.
GN 19.5.
OS Bacteriophage T7.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;
OC T7-like viruses.
OC NCBI_TaxID=10760;
OX [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83241725; PubMed=6864790;
RA Dunn J.J., Studier F.W.;
RT "Complete nucleotide sequence of bacteriophage T7 DNA and the
RT locations of T7 genetic elements.";
CC J. Mol. Biol. 168:477-535(1983).
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, V01146; CAA24443.1; -
CC DR PIR: A04429; O9BPE7.
DR PIR: S42341; S42341.
KM Hypothetical protein.
SQ SEQUENCE 49 AA; 5434 MW; FB1C9240BA465342 CRC64;
-----
Query Match 43.1%; Score 25; DB 1; Length 49;
Best Local Similarity 44.4%; Pred. No. 1.5e+02;
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
-----
QY 4 LRRYILRY 12
DB 9 LRRHYTYR 17
-----
RESULT 5
YL9A_LACCU STANDARD; PRT; 51 AA.
AC P35923;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 6.1 kDa protein in CURA 3' region (ORF8).
OS Lactobacillus curvatus.
OG Plasmid 60 kb.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OC NCBI_TaxID=28038;
OX [1]
RP SEQUENCE FROM N.A.
RX STRAIN=LTH1174;
RX MEDLINE=94058500; PubMed=7694558;

```

```

RA Tichaczek P.S., Vogel R.P., Hammes W.P.;
RT "Cloning and sequencing of curA encoding curvacin A, the bacteriocin
RL produced by Lactobacillus curvatus LTH1174."
RL Arch. Microbiol. 160:279-283(1993).
CC -1- FUNCTION: HAS SOME FUNCTION IN CURVACIN A BIOSYNTHESIS, TRANSPORT
CC OR RESISTANCE
KW Plasmid; Hypothetical protein.
SQ SEQUENCE 51 AA; 6142 MW; E77506B2076B7CPE CRC64;

Query Match 43.1%; Score 25; DB 1; Length 51;
Best Local Similarity 45.5%; Pred. No. 1.5e+02;
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 2 IALRYRILLY 12
DB 12 LALQLIILRF 22

RESULT 6
VSH_MUMP2 STANDARD; PRT; 57 AA.
AC P28084;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Small hydrophobic protein.
GN SH.
OS Mumps virus (strain Edingburgh 2), and
OS Mumps virus (strain Edingburgh 6).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Rubulaviruses;
OC NCBI_TaxID=11162, 11164;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93168036; PubMed=8435047;
RA Yeo R.P., Afzal M.A., Forsey T., Rima B.K.;
RT "Identification of a new mumps virus lineage by nucleotide sequence
analysis of the SH gene of ten different strains."
RL Arch. Virol. 128:371-377(1993).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X63711; CAA45244.1; -
DR PIR; S19867; SHNZE4.
DR InterPro; IPR001477; SH.
DR Pfam; PF01445; SH; 1.
DR ProDom; PD001504; SH; 1.
KW Transmembrane.
FT TRANSMEM 8 32 POTENTIAL.
SQ SEQUENCE 57 AA; 6894 MW; 49121303FE0DED94 CRC64;

Query Match 43.1%; Score 25; DB 1; Length 51;
Best Local Similarity 45.5%; Pred. No. 1.7e+02;
Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 IALRYRILLY 12
DB 15 LMLLYRIITLY 25

RESULT 7
VSH_MUMP4 STANDARD; PRT; 57 AA.
AC P28085;
DT 01-AUG-1992 (Rel. 23, Created)

```

```

DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Small hydrophobic protein.
GN SH.
OS Mumps virus (strain Edingburgh 4).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Rubulaviruses;
OC NCBI_TaxID=11163;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93168036; PubMed=8435047;
RA Yeo R.P., Afzal M.A., Forsey T., Rima B.K.;
RT "Identification of a new mumps virus lineage by nucleotide sequence
analysis of the SH gene of ten different strains."
RL Arch. Virol. 128:371-377(1993).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X63710; CAA45243.1; -
DR PIR; S19867; SHNZE4.
DR InterPro; IPR001477; SH.
DR Pfam; PF01445; SH; 1.
DR ProDom; PD001504; SH; 1.
KW Transmembrane.
FT TRANSMEM 8 32 POTENTIAL.
SQ SEQUENCE 57 AA; 6852 MW; 191F1E03FE0DE57F CRC64;

Query Match 43.1%; Score 25; DB 1; Length 57;
Best Local Similarity 45.5%; Pred. No. 1.7e+02;
Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 IALRYRILLY 12
DB 15 LMLLYRIITLY 25

RESULT 8
VSH_MUMPB STANDARD; PRT; 57 AA.
AC P28082;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Small hydrophobic protein.
GN SH.
OS Mumps virus (strain Belfast).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Rubulaviruses;
OC NCBI_TaxID=11166;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93168036; PubMed=8435047;
RA Yeo R.P., Afzal M.A., Forsey T., Rima B.K.;
RT "Identification of a new mumps virus lineage by nucleotide sequence
analysis of the SH gene of ten different strains."
RL Arch. Virol. 128:371-377(1993).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X63709; CAA45242.1; -
DR PIR; S19866; SHN2BF.

```

```

DR InterPro; IPR001477; SH.
DR Pfam; PF01445; SH; 1.
DR ProDom; PD001504; SH; 1.
KW Transmembrane.
FT TRANSMEM 8 32 POTENTIAL.
SQ SEQUENCE 57 AA; 6814 MW; B8F2F917B4ABEF3C CRC64;

Query Match 43.1%; Score 25; DB 1; Length 57;
Best Local Similarity 45.5%; Pred. No. 1.7e+02;
Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 2 IALRYILRLY 12
DB 15 LILLYRIITLY 25

RESULT 9
VSH_MUMPR STANDARD; PRT; 57 AA.
ID VSH_MUMPR STANDARD; PRT; 57 AA.
AC P28087;
DT 01-AUG-1992 (Rel. 23; Created)
DT 01-AUG-1992 (Rel. 23; Last sequence update)
DT 01-NOV-1997 (Rel. 35; Last annotation update)
DE Small hydrophobic protein.
GN SH.
OS Mumps virus (strain RW).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Rubulavirus.
OX NCBI_Taxid=11172;
RA [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93168036; PubMed=8435047;
RA Yeo R.P.; Afzal M.A.; Forsey T.; Rima B.K.;
RT "Identification of a new mumps virus lineage by nucleotide sequence
RT analysis of the SH gene of ten different strains.";
RL Arch. Virol. 128:371-377(1993).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X63708; CAA45241.1; -.
DR PIR; S24827; S24827.
DR InterPro; IPR001477; SH.
DR Pfam; PF01445; SH; 1.
DR ProDom; PD001504; SH; 1.
KW Transmembrane.
FT TRANSMEM 8 32 POTENTIAL.
SQ SEQUENCE 57 AA; 6892 MW; B8FASB0945BC04D3 CRC64;

Query Match 43.1%; Score 25; DB 1; Length 57;
Best Local Similarity 45.5%; Pred. No. 1.7e+02;
Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 2 IALRYILRLY 12
DB 15 LILLYRIITLY 25

RESULT 10
V560_HELPY STANDARD; PRT; 26 AA.
ID V560_HELPY STANDARD; PRT; 26 AA.
AC O52585;
DT 16-OCT-2001 (Rel. 40; Created)
DT 16-OCT-2001 (Rel. 40; Last sequence update)
DT 16-OCT-2001 (Rel. 40; Last annotation update)
DE Hypothetical protein HP0560/JHP0507.
GN HP0560 OR JHP0507.
OS Helicobacter pylori (Campylobacter pylori), and

```

```

OS Helicobacter pylori J99 (Campylobacter pylori J99).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_Taxid=210, 85963;
(1)
RP SEQUENCE FROM N.A.
RX STRAIN=26695 / ATCC 700392;
RX MEDLINE=97394467; PubMed=9252185;
RA Tomasz J.-F., White O., Keriavage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA Loftus B., Richardson D., Dodson R., Khalaf H.G., Glodek A.,
RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.D., Melman J.M., Fujii C., Bowman C., Watney L., Wallin E.,
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter
RT pylori.";
RL Nature 388:539-547(1997).
RN (2)
RP SEQUENCE FROM N.A.
RX STRAIN=J99;
RX MEDLINE=99120557; PubMed=9923682;
RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA Smith D.R., Noonan B., Guid B.C., deJonge B.L., Carmel G.,
RA Tibbino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Yovis G.F.,
RA Trust T.J.;
RT "Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen Helicobacter pylori.";
RL Nature 397:176-180(1999).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000570; AAD07633.1; -.
DR EMBL; AE001484; AAD06083.1; -.
DR TIGR; HP0560; -.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 3 23 POTENTIAL.
SQ SEQUENCE 26 AA; 3169 MW; 34FE239BE3139A1 CRC64;

Query Match 41.4%; Score 24; DB 1; Length 26;
Best Local Similarity 75.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 IALRYIL 9
DB 13 IYLRLYL 20

RESULT 11
VPU_HVIA2 STANDARD; PRT; 38 AA.
ID VPU_HVIA2 STANDARD; PRT; 38 AA.
AC P05549;
DT 01-NOV-1988 (Rel. 09; Created)
DT 01-NOV-1988 (Rel. 09; Last sequence update)
DT 01-NOV-1997 (Rel. 35; Last annotation update)
DE VPU protein (U ORF protein).
GN VPU
OS Human immunodeficiency virus type 1 (ARV2/SF2 isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_Taxid=11685;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=85090453; PubMed=2578227;
RA Sanchez-Pescador R., Power M.D., Barr P.J., Steimer K.S.,

```

RA Stempjien M.M., Brown-Shimer S.L., Gee W.W., Renard A., Randolph A.,
 RT Levy J.A., Dina D., Luciw P.A.;
 "Nucleotide sequence and expression of an AIDS-associated retrovirus
 (ARV-2).",
 RL Science 227:484-492(1985).
 CC -1- FUNCTION: ACTS IN THE DEGRADATION OF CD4 IN THE ENDOPLASMIC
 CC RETICULUM AND IN THE ENHANCEMENT OF VIRION RELEASE FROM THE PLASMA
 CC MEMBRANE OF INFECTED CELLS.
 CC -1- SUBCELLULAR LOCATION: Membrane-bound.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; K02007; AAB5881.1; ..
 CC HIV; K02007; VPOSSP2.
 CC InterPro; IPR002094; Vpu.
 CC Pfam; PF00558; Vpu; 1.
 CC Transmembrane; AIDS.
 CC TRANSMEM 8 28
 CC SEQUENCE 38 AA; 4390 MW; 7F6431F505BEB6C7 CRC64;
 CC -----
 CC POTENTIAL.
 CC
 CC Query Match 41.4%; Score 24; DB 1; Length 38;
 CC Best Local Similarity 40.0%; Pred. No. 1.7e+02;
 CC Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 CC
 CC QY 2 IALRYRILLR 11
 CC : : : : :
 CC Db 26 VLIEYRKILR 35
 CC -----
 CC RESULT 12
 CC Y195_BPT3 STANDARD; PRT; 49 AA.
 CC ID Y195_BPT3
 CC AC P10306;
 CC DT 01-MAR-1989 (Rel. 10, Created)
 CC DT 01-MAR-1989 (Rel. 10, Last sequence update)
 CC DT 01-APR-1990 (Rel. 14, Last annotation update)
 CC DE Hypothetical gene 19.5 protein.
 CC GN 19.5.
 CC OS Bacteriophage T3.
 CC OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;
 CC OT T7-like viruses.
 CC OX NCBI_TaxID=10759;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RX MEDLINE=86209937; PubMed=3010556;
 CC RA Yamada M., Fujisawa H., Kato H., Hamada K., Minagawa T.;
 RT "Cloning and sequencing of the genetic right end of bacteriophage T3
 RT DNA".
 RT Virology 151:350-361(1986).
 RL [2]
 RN ERRATUM.
 RP Yamada M., Fujisawa H., Kato H., Hamada K., Minagawa T.;
 RL Virology 154:246-246(1986).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M14784; AAA32531.1; ..
 CC PIR; I23476; O9BPT3.
 CC KW Hypothetical protein.
 CC SEQUENCE 49 AA; 5484 MW; 266E0EF6CC4B52AF CRC64;
 CC -----

Query Match 41.4%; Score 24; DB 1; Length 49;
 Best Local Similarity 44.4%; Pred. No. 2.3e+02;
 Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 QY 4 LRYRILLRY 12
 Db 9 LRHRVTWRF 17

 RESULT 13
 RL32 BORBU STANDARD; PRT; 59 AA.
 ID RL32 BORBU
 AC O51646;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE 50S ribosomal protein L32.
 GN RPFM OR BB0703.
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_TaxID=139;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 35210 / B31;
 RX MEDLINE=98065943; PubMed=9403685;
 RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
 RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
 RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D., Hanson M.,
 RA Peterson J., Karpavage A.R., Quackenbush J., Salzberg S., Weidman J.,
 RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
 RA Utterback T., Watthey L., McDonald L., Artiach P., Bowman C.,
 RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
 RA Smith H.O., Venter J.C.;
 RT "Genomic sequence of a Lyme disease spirochete, Borrelia
 RT burgdorferi".
 RT Nature 390:580-586(1997).
 RL CC -1- SIMILARITY: BELONGS TO THE L32P FAMILY OF RIBOSOMAL PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AE001170; AAC67042.1; ..
 CC DR TIGR; BB0703; ..
 CC DR InterPro; IPR002677; Ribosomal L32p.
 CC DR Pfam; PF01783; Ribosomal L32p; 1.
 CC DR TIGRPFAM; TIGR01031; rplMf bact; 1.
 CC KW Ribosomal protein; Complete proteome.
 CC FT INIT_MET 0 BY SIMILARITY.
 CC SQ SEQUENCE 59 AA; 6945 MW; 958EC3FE0F8D87E CRC64;
 CC -----
 CC Query Match 41.4%; Score 24; DB 1; Length 59;
 CC Best Local Similarity 40.0%; Pred. No. 2.8e+02;
 CC Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
 CC
 CC QY 2 IALRYRILLR 11
 CC : : : : :
 CC Db 36 LGVRRHICLK 45
 CC -----
 CC RESULT 14
 CC YC4_ODOSI
 ID YC4_ODOSI STANDARD; PRT; 29 AA.
 AC P49830;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical 3.5 kDa protein in YCF33-TRN intergenic region (ORF29A).
 OS Odontella sinensis (Marine centric diatom).
 CC -----

CC Chloroplast.
 CC Eukaryota; stramenopiles; Bacillariophyta; Coscinodiscophyceae;
 CC Bidulphiophycidae; Eupodiscaceae; Eupodiscaceae; Odontella.
 CC NCBI_TaxID=2839;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kowalik K.V., Stroebe B., Schaffran I., Kroth-Pancic P., Freier U.,
 RT "The chloroplast genome of a chlorophyll a+c-containing alga,
 RL Plant Mol. Biol. Rep. 13:336-342(1995).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; Z67753; CAA91683.1; -
 DR Chloroplast; Hypothetical protein.
 KW SEQUENCE 29 AA; 3481 MW; 4960F2AC972ACBA2 CRC64;
 SQ

Query Match 39.7%; Score 23; DB 1; Length 29;
 Best Local Similarity 80.0%; Pred. No. 2e+02; 0; Gaps 0;
 Matches 4; Conservative 1; Mismatches 0; Indels 0;

OY 8 ILTRY 12
 :|||
 Db 7 VLTRY 11

RESULT 15
 ID LHG_RHOVI STANDARD; PRT; 36 AA.
 AC P04126;
 DT 20-MAR-1987 (Rel. 04, Created)
 DT 20-MAR-1987 (Rel. 04, Last sequence update)
 DT 01-NOV-1991 (Rel. 20, Last annotation update)
 DE Light-harvesting protein B-1015, gamma chain.
 OS Rhodospseudomonas viridis.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Hyphomicrobium group; Blastochloris.
 CC NCBI_TaxID=1079;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=85225948; PubMed=3890891;
 RA Brunisholz R.A., Jay F., Suter F., Zuber H.;
 RT "The light-harvesting polypeptides of Rhodospseudomonas viridis. The
 RT complete amino-acid sequences of B1015-alpha, B1015-beta and B1015-
 RT gamma.";
 RL Biol. Chem. Hoppe-Seyler 366:87-98(1985).
 CC -!- FUNCTION: ONE OF THE COMPONENTS OF THE BACTERIOCHLOROPHYLL-PROTEIN
 CC COMPLEX IN THE CHROMATOPHORE MEMBRANE.
 CC PIR: A03455; LBRGV.
 DR Antenna complex; Light-harvesting polypeptide.
 KW VARIANT 34 34 T -> V (IN 33% OF THE MOLECULES).
 FT SEQUENCE 36 AA; 4006 MW; EC6F189A0F69F36A CRC64;
 SQ

Query Match 39.7%; Score 23; DB 1; Length 36;
 Best Local Similarity 83.3%; Pred. No. 2.6e+02; 0; Gaps 0;
 Matches 5; Conservative 0; Mismatches 1; Indels 0;

OY 2 IALRY 7
 :|||
 Db 31 IALTYR 36

Search completed: December 19, 2002, 16:19:45
 Job time: 5.47059 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 19, 2002, 16:16:02 ; Search time 10.4118 Seconds
(without alignments)
110.799 Million cell updates/sec

Title: US-08-653-294C-41

Perfect score: 58

Sequence: 1 RIALRYRILLRY 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 15911

Minimum DB seq length: 0
Maximum DB seq length: 60

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	29	50.0	44	4	IMBP8
2	29	50.0	44	2	B82581
3	28	48.3	53	2	B83853
4	28	48.3	58	2	C97788
5	28	48.3	60	2	T03816
6	27	46.6	50	2	T13307
7	27	46.6	52	2	F96947
8	26	44.8	15	2	C48401
9	25	43.1	29	2	B81006
10	25	43.1	46	1	S65884
11	25	43.1	46	1	S65591
12	25	43.1	48	2	T07202
13	25	43.1	49	1	Q9BPE7
14	25	43.1	50	2	C72257
15	25	43.1	51	2	S72280
16	25	43.1	57	1	SHNZBF
17	25	43.1	57	1	SHNZE4
18	25	43.1	57	1	SHNZE2
19	25	43.1	57	2	S24827
20	25	43.1	58	2	B82816
21	25	43.1	58	2	B86774
22	25	43.1	58	2	B97141
23	25	43.1	59	2	B97654
24	24	41.4	26	2	H64589
25	24	41.4	27	4	I52725
26	24	41.4	49	1	Q9BPT3
27	24	41.4	56	2	C90017
28	24	41.4	56	2	A12134
29	24	41.4	57	2	JQ2368

30 24 41.4 58 2 A86702 hypothetical prote
31 24 41.4 58 2 C82527 hypothetical prote
32 24 41.4 58 2 C95331 hypothetical prote
33 24 41.4 60 1 F70187 ribosomal protei
34 23 39.7 29 2 S78310 hypothetical prote
35 23 39.7 30 2 A28562 glutathione transf
36 23 39.7 33 2 S17164 glutathione transf
37 23 39.7 34 2 G81597 hypothetical prote
38 23 39.7 36 1 LRRFGV light-harvesting p
39 23 39.7 37 2 C64219 ribosomal protei
40 23 39.7 40 2 A64530 hypothetical prote
41 23 39.7 43 2 A81532 hypothetical prote
42 23 39.7 46 1 SWLQDA diuretic hormone -
43 23 39.7 46 2 A48542 CRF-related diuret
44 23 39.7 52 2 D84640 hypothetical prote
45 23 39.7 55 2 S33915 hypothetical prote

ALIGNMENTS

RESULT 1

IMBP8

hypothetical immunity region protein 8 - Bacillus phage phi-105

C:Species: Bacillus phage phi-105

A:Note: host Bacillus subtilis

C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999

C:Accession: G24521

R:Cully, D.F.; Garro, A.J.

Gene 38, 153-164, 1985

A:Title: Nucleotide sequence of the immunity region of Bacillus subtilis bacteriophage phi-105

A:Reference number: A91535, MUID:86056972; PMID:3934047

A:Accession: G24521

A:Molecule type: DNA

A:Residues: 1-44 <CUL>

A:Cross-references: GB:M1920; NID:g215477

C:Comment: This is the hypothetical translation of a sequence that was not reported as a

Query Match 50.0%; Score 29; DB 4; Length 44;

Best Local Similarity 66.7%; Pred. No. 71;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RIALRYRILL 9

Db 15 RLALKVRIL 23

RESULT 2

B82581

hypothetical protein XF2242 [imported] - Xylella fastidiosa (strain 9a5c)

C:Species: Xylella fastidiosa

C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000

C:Accession: B82581

R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing

Nature 406, 151-157, 2000

A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.

A:Reference number: A82515; MUID:20365717; PMID:10910347

A:Note: for a complete list of authors see reference number A59328 below

A:Accession: B82581

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-54 <SIM>

A:Cross-references: GB:AE003849; NID:g9107394; PIDN:AAF85041.1; GSPDB:GN001

A:Experimental source: strain 9a5c

R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, F.

as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.

submitted to GenBank, June 2000

A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohman, J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.B.; Laigret, M.A.; Madeira, A.M.B.N.; Madeira, L.M.F.; Marino, C.L.; Marques, M.V.; Martins, F.

A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.F.

Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A.; Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tsunako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF2242

Query Match 50.0%; Score 29; DB 2; Length 54;
Best Local Similarity 45.5%; Pred. No. 88;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
Qy 2 LRLRYRLRLY 12
Db 29 LRLNHAIRLY 39

RESULT 3
B83853
hypothetical protein BH1626 [imported] - *Bacillus halodurans* (strain C-125)
C:Species: *Bacillus halodurans*
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: B83853
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Maenu, N.; Fuji, F.; Hirz
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and
A:Reference number: A63650; MUID:20512582; PMID:11058132
A:Accession: B83853
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-53 <STO>
A:Cross-references: GB:AP001512; GB:BA000004; NID:g10174030; PIDN:BA05345.1; GSPDB:GN00
A:Experimental source: strain C-125
C:Genetics:
A:Gene: BH1626

Query Match 48.3%; Score 28; DB 2; Length 53;
Best Local Similarity 57.1%; Pred. No. 1.3e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
Qy 5 RYRILRL 11
Db 5 RYRIVIR 11

RESULT 4
C97788
hypothetical protein RC0707 [imported] - *Rickettsia conorii* (strain Malish 7)
C:Species: *Rickettsia conorii*
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 30-Sep-2001
C:Accession: C97788
R:Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; Ro
Science 293, 2093-2098, 2001
A:Title: Mechanisms of Evolution in *Rickettsia conorii* and *Rickettsia prowazekii*.
A:Reference number: A97700; MUID:21442074; PMID:11557893
A:Accession: C97788
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-58 <KUR>
A:Cross-references: GB:AB006914; PIDN:AAL03245.1; PID:g15619799; GSPDB:GN00173
C:Genetics:
A:Gene: RC0707

Query Match 48.3%; Score 28; DB 2; Length 58;
Best Local Similarity 58.3%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 1 RIALRYRLRLY 12
Db 43 RRAMRYMFLLY 54

RESULT 5

T03816
hypothetical protein (clone NG2) - common tobacco
C:Species: *Nicotiana tabacum* (common tobacco)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000
C:Accession: T03816
R:Karrer, E.E.; Beachy, R.N.; Holt, C.A.
Plant Mol. Biol. 36, 681-690, 1998
A:Title: Cloning of tobacco genes that elicit the hypersensitive response.
A:Reference number: Z15096; MUID:98187248; PMID:9526500
A:Accession: T03816
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-60 <KAR>
A:Cross-references: EMBL:U66272; NID:g1762950; PIDN:AAC49978.1; PID:g1762951
A:Experimental source: clone NG2

Query Match 48.3%; Score 28; DB 2; Length 60;
Best Local Similarity 44.4%; Pred. No. 1.5e+02;
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
Qy 4 LRYRLRLY 12
Db 18 RYKYLWLY 26

RESULT 6
T13307
hypothetical protein 18 - *Streptococcus phage phi-O1205*
C:Species: *Streptococcus phage phi-O1205*
C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 24-Sep-1999
C:Accession: T13307
R:Stanley, E.; Fitzgerald, G.F.; Le Marrec, C.; Payard, B.; van Sinderen, D.
Microbiology 143, 3417-3429, 1997
A:Title: Sequence analysis and characterization of phi O1205, a temperate bacteriophage
A:Reference number: Z17654; MUID:98048466; PMID:9387220
A:Accession: T13307
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-50 <STA>
A:Cross-references: EMBL:U86974; NID:g2444080; PID:g2444098; PIDN:AACT9534.1
A:Experimental source: host *Streptococcus thermophilus* strain CNRZ1205

Query Match 46.6%; Score 27; DB 2; Length 50;
Best Local Similarity 33.3%; Pred. No. 2e+02;
Matches 3; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
Qy 4 LRYRLRLY 12
Db 1 MRYKIVLY 9

RESULT 7
F96947
hypothetical protein CAC0389 [imported] - *Clostridium acetobutylicum*
C:Species: *Clostridium acetobutylicum*
C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
C:Accession: F96947
R:Noelling, J.; Breton, G.; Ometchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
J. Dairy Sci. 82, 4823-4838, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium *Clo*
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: F96947
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-52 <KUR>
A:Cross-references: GB:AF001437; PIDN:AAK78369.1; PID:g15023240; GSPDB:GN00168
A:Experimental source: *Clostridium acetobutylicum* ATCC824
C:Genetics:
A:Gene: CAC0389

Query Match 46.6%; Score 27; DB 2; Length 52;
Best Local Similarity 54.5%; Pred. No. 2.1e+02;

Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 RIALRYRILLR 11
| | | | |
Db 11 RCANSYRFLR 21

RESULT 8
C48401
ribosomal protein L1 - Thermus aquaticus (fragment)
C/Species: Thermus aquaticus
C/Date: 01-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 18-Nov-1994
C/Accession: C48401
R/Gardner, M.B.; Agalarov, S.C.; Eliseikina, I.A.; Fomenkova, N.P.; Nikonov, S.V.; Sedeln
Biochimie 74, 327-336, 1992
A/Title: Ribosomal proteins from Thermus thermophilus for structural investigations.
A/Reference number: A48401; MUID:92345325; PMID:1637860
A/Accession: C48401
A/Status: preliminary
A/Molecule type: protein
A/Residues: 1-15 <GAR>
A/Note: sequence extracted from NCBI backbone (NCBIP:109932)

Query Match 44.8%; Score 26; DB 2; Length 15;
Best Local Similarity 62.5%; Pred. No. 90;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 5 RYRILLRY 12
| | | | |
Db 6 RYRALLXF 13

RESULT 9
B81006
hypothetical protein NMB2097 [imported] - Neisseria meningitidis (strain MC58 serogroup
C/Species: Neisseria meningitidis
C/Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C/Accession: B81006
R/Tetzelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
xi, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A/Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve
A/Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A/Reference number: B81000; MUID:20175755; PMID:10710307
A/Accession: B81006
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-29 <TET>
A/Cross-references: GB:AE002559; GB:AE002098; NID:g7227357; PIDN:AAF42414.1; PID:g722736
A/Experimental source: serogroup B, strain MC58
C/Genetics:
A/Gene: NMB2097

Query Match 43.1%; Score 25; DB 2; Length 29;
Best Local Similarity 50.0%; Pred. No. 2.8e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 5 RYRILLRY 12
| | | | |
Db 19 RPSLLRFP 26

RESULT 10
S66584
hypothetical protein 45 - phase 82
C/Species: phase 82
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C/Accession: S66584
R/Mahdi, A.A.; Sharples, G.J.; Mandal, T.N.; Lloyd, R.G.
J. Mol. Biol. 257, 561-573, 1996
A/Title: Holliday junction resolvases encoded by homologous rnaA genes in Escherichia co
A/Reference number: S66579; MUID:96196428; PMID:8648624

A/Accession: S66584
A/Molecule type: DNA
A/Residues: 1-46 <MAH>
A/Cross-references: EMBL:X92588; NID:gi051111; PIDN:CAA63331.1; PID:gi051117
C/Superfamily: phase 82 hypothetical protein 45

Query Match 43.1%; Score 25; DB 1; Length 46;
Best Local Similarity 71.4%; Pred. No. 4.5e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 YRILLRY 12
| | | | |
Db 31 YRILKNY 37

RESULT 11
S66591
hypothetical protein 45 - Escherichia coli cryptic lambdaoid prophage DLP12
C/Species: Escherichia coli
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C/Accession: S66591
R/Mahdi, A.A.; Sharples, G.J.; Mandal, T.N.; Lloyd, R.G.
J. Mol. Biol. 257, 561-573, 1996
A/Title: Holliday junction resolvases encoded by homologous rnaA genes in Escherichia co
A/Reference number: S66579; MUID:96196428; PMID:8648624
A/Accession: S66591
A/Molecule type: DNA
A/Residues: 1-46 <MAH>
A/Cross-references: EMBL:X92587; NID:gi0511136; PIDN:CAA63322.1; PID:gi051142
C/Genetics:
A/Map position: 12 min
A/Genome: prophage
C/Superfamily: phase 82 hypothetical protein 45

Query Match 43.1%; Score 25; DB 1; Length 46;
Best Local Similarity 71.4%; Pred. No. 4.5e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 YRILLRY 12
| | | | |
Db 31 YRILKNY 37

RESULT 12
T07202
hypothetical protein 48a - Chlorella vulgaris chloroplast
C/Species: chloroplast Chlorella vulgaris
C/Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 21-Jul-2000
C/Accession: T07202
R/Wakasugi, T.; Nagai, T.; Kapoor, M.; Sugita, M.; Ito, M.; Ito, S.; Tsudzuki, J.; Nakai
Proc. Natl. Acad. Sci. U.S.A. 94, 5967-5972, 1997
A/Title: Complete nucleotide sequence of the chloroplast genome from the green alga Chl
A/Reference number: Z15985; MUID:97303241; PMID:9159184
A/Accession: T07202
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-48 <WAK>
A/Cross-references: EMBL:AB001684; NID:g2224352; PIDN:BAA57849.1; PID:g2224365
C/Genetics:
A/Genome: chloroplast
C/Keywords: chloroplast

Query Match 43.1%; Score 25; DB 2; Length 48;
Best Local Similarity 62.5%; Pred. No. 4.6e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 5 RYRILLRY 12
| | | | |
Db 9 RYFILLRY 16

RESULT 13
Q9BPE7

gene 19.5 protein - phage T7
 C:Species: phage T7
 C:Date: 13-Jun-1983 #sequence_revision 13-Jun-1983 #text_change 23-Jul-1999
 C:Accession: A04429; S42341
 R:Dunn, J.J.; Thompson, K.
 submitted to the Nucleic Acid Sequence Database, September 1982
 A:Reference number: A94615
 A:Accession: A04429
 A:Molecule type: DNA
 A:Residues: 1-49 <DUN>
 R:Dunn, J.J.; Studier, F.W.
 J. Mol. Biol. 166, 477-535, 1983
 A:Title: Complete nucleotide sequence of bacteriophage T7 DNA and the locations of T7 gene
 A:Reference number: S42283; MUID:83241725; PMID:6864790
 A:Accession: S42341
 A:Molecule type: DNA
 A:Residues: 1-49 <DUN>
 A:Cross-references: EMBL:V01146; NID:g431187; PIDN:CAA24443.1; PID:g15619
 C:Genetics:
 A:Gene: 19.5
 A:Map position: 98.63-98.99
 C:Superfamily: phage T7 gene 19.5 protein

Query Match 43.1%; Score 25; DB 1; Length 49;
 Best Local Similarity 44.4%; Pred. No. 4.7e+02;
 Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 4 LRYRILRY 12
 ||: ||: ||:
 Db 9 LRRRTYRF 17

RESULT 14
 C72257
 hypothetical protein - Thermotoga maritima (strain MSB8)
 C:Species: Thermotoga maritima
 C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
 C:Accession: C72257
 R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwin, M.L.; Dodson, R.J.; Haft, D.H.; Hickey
 Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
 C.M.
 Nature 399, 323-329, 1999
 A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
 A:Reference number: A72200; MUID:99287316; PMID:10360571
 A:Accession: C72257
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-50 <ARN>
 A:Cross-references: GB:AB001793; GB:AB000512; NID:g4981963; PIDN:AA036473.1; PID:g498196
 A:Experimental source: strain MSB8
 C:Genetics:
 A:Gene: TML402

Query Match 43.1%; Score 25; DB 2; Length 50;
 Best Local Similarity 50.0%; Pred. No. 4.8e+02;
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 2 IALRYILR 11
 : ||: ||:
 Db 22 IALNVRIMR 31

RESULT 15
 S72280
 hypothetical protein 51 - Plasmodium falciparum plastid
 C:Species: plasmodium falciparum
 C:Date: 14-Apr-1998 #sequence_revision 24-Apr-1998 #text_change 08-Oct-1999
 C:Accession: S72280
 R:Wilson, R.J.M.; Denny, P.W.; Preiser, P.R.; Rangachari, K.; Roberts, K.; Roy, A.; Whyt
 J. Mol. Biol. 261, 155-172, 1996
 A:Title: Complete gene map of the plastid-like DNA of the malaria parasite Plasmodium fa
 A:Reference number: S72277; MUID:96346169; PMID:8757284
 A:Accession: S72280

A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-51 <MIL>
 A:Cross-references: EMBL:X95275; NID:g1171583; PIDN:CAA64570.1; PID:e220241; PID:g117158
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1996
 C:Genetics:
 A:Genome: plastid
 A:Note: this apparently degenerate plastid is referred to as the apicoplast
 C:Keywords: plastid

Query Match 43.1%; Score 25; DB 2; Length 51;
 Best Local Similarity 50.0%; Pred. No. 4.9e+02;
 Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 5 RYRILRY 12
 : ||: ||:
 Db 17 KYKLLRY 24

Search completed: December 19, 2002, 16:22:59
 Job time : 11.418 secs

XX	OS	Hom sapiens.
XX	XX	
XX	PN	WO9958660-Al.
XX	XX	
XX	PD	18-NOV-1999.
XX	XX	
XX	PF	06-MAY-1999;
XX	XX	
PR	PR	12-MAY-1998;
PR	PR	12-MAY-1999;
PR	PR	12-MAY-1998;
PR	PR	12-MAY-1999;
PR	PR	18-MAY-1999;
PR	PR	18-MAY-1999;
PR	PR	18-MAY-1998;

```

PR 18-MAY-1998; 98US-0085922.
PR 18-MAY-1998; 98US-0085923.
PR 18-MAY-1998; 98US-0085924.
PR 18-MAY-1998; 98US-0085925.
PR 18-MAY-1998; 98US-0085926.
PR 18-MAY-1998; 98US-0085927.
XX
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ruben SM, Florence K, Ni J, Rosen CA, Carter KC, Moore PA;
PI Olsen HS, Shi Y, Young PE, Wei F, Brewer LA, Soppet DR;
PI Lafleur DW, Edress GA, Ebner R;
XX
XX WPI: 2000-062296/05.
XX
XX New isolated human genes and the secreted polypeptides they encode,
XX useful for diagnosis and treatment of e.g. cancers, neurological
XX disorders, immune diseases, inflammation or blood disorders -
XX
XX
XX Disclosure: Page 450; 475pp; English.
XX
XX AA65250 to AA65350 represent 97 isolated human secreted protein genes.
XX AA76124 to AA76223 are the secreted proteins encoded by the 97 human
XX genes. This sequence represents a fragment of one of the human secreted
XX proteins. The genes and their corresponding secreted polypeptides are
XX useful for preventing, treating or ameliorating medical conditions,
XX e.g. by protein or gene therapy. Also pathological conditions can be
XX diagnosed by determining the amount of the new polypeptides in a sample
XX or by determining the presence of mutations in the new genes. Specific
XX uses are described for each of the 97 genes, based on which tissues they
XX are most highly expressed in, and include developing products for the
XX diagnosis or treatment of cancer, tumours, developmental abnormalities
XX and foetal deficiencies, blood disorders, diseases of the immune system,
XX autoimmune diseases, inflammation, allergies, Alzheimer's and cognitive
XX disorders, schizophrenia, arthritis, asthma, psoriasis, sepsis, skin
XX disorders, atherosclerosis, diabetes, cardiovascular disorders, kidney
XX disorders, digestive/endocrine disorders, infections and AIDS. The
XX polypeptides are also useful for identifying their binding partners.
XX The sequences shown in AA76224 to AA76424 represent fragments of the
XX secreted proteins.
XX
XX
XX Sequence 38 AA;
XX
XX Query Match 67.2%; Score 39; DB 21; Length 38;
XX Best Local Similarity 70.0%; Pred. No. 2.5;
XX Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
XX
XX 3 AARYRLAIR 12
XX |:|:|:|:|
XX Db 9 AIRYREAVR 18
XX
XX
XX RESULT 2
XX ID AAUS9703 standard; Protein, 57 AA.
XX
XX AAUS9703;
XX
XX 27-FEB-2002 (first entry)
XX
XX Propionibacterium acnes immunogenic protein #20599.
XX
XX SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis;
XX uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
XX inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
XX dermatological; osteopathic; neuroprotectant.
XX
XX Propionibacterium acnes.
XX
XX WO200181581-A2.
XX
XX 01-NOV-2001.
XX

```

```

PF 20-APR-2001; 2001WO-US12865.
XX
XX 21-APR-2000; 2000US-199047P.
PR 02-JUN-2000; 2000US-208841P.
PR 07-JUL-2000; 2000US-216747P.
XX
XX (CORI-) CORIXA CORP.
XX
XX Skeiky YAM, Persing DH, Mitcham JT, Wang SS, Bhactia A;
PI L'maisonmeuve J, Zhang Y, Jen S, Carter D;
DR WPI; 2001-616774/71.
XX
XX N-PSDB; AAS59605.
XX
XX Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris -
XX
XX
XX Example 1; SEQ ID NO 20898; 1069pp; English.
XX
XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPPO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX
XX Sequence 57 AA;
XX
XX Query Match 67.2%; Score 39; DB 22; Length 57;
XX Best Local Similarity 58.3%; Pred. No. 3.9;
XX Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
XX
XX 1 RLAFFYRLAIR 12
XX |:|:|:|:|:|
XX
XX Db 18 RCSLRYRLSMR 29
XX
XX
XX RESULT 3
XX ID AAY72482 standard; peptide; 10 AA.
XX
XX AAY72482;
XX
XX 24-APR-2001 (first entry)
XX
XX Immunosuppressive peptide, P15, to prevent allograft rejection.
XX
XX Immunosuppressive; allograft rejection; topological parameter;
XX physico-chemical parameter; in silico screening; pharmaceutical;
XX cosmetic; agrochemical; biomaterial; veterinary application.
XX
XX Unidentified.
XX
XX OS
XX
XX PN WO200079263-A2.
XX
XX PD 28-DEC-2000.
XX
XX 15-MAY-2000; 2000WO-EP04338.
XX

```


CC The method is useful for transplanting a donor mammalian organ or cells
 CC to a mammalian recipient. The organ or cells are combined with the
 CC compound prior to implanting in the mammalian recipient, or the compound
 CC may be administered to the mammalian recipient in a period extending
 CC from prior to implanting the donor organ or cells.

SO Sequence 14 AA;

Query Match 62.1%; Score 36; DB 22; Length 14;
 Best Local Similarity 100.0%; Pred. No. 3.1;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RIALRY 7
 DB 8 RIALRY 14

RESULT 6
 AAB59417

ID AAB59417 standard; Peptide; 14 AA.

XX AAB59417;

DT 22-MAR-2001 (first entry)

DE Human Class I HLA-B alpha1-domain-derived peptide #17.

XX Human; human leukocyte antigen-B; HLA-B; immunosuppressive;

KW cytotoxic T cell activation inhibition; cell transplantation;

KW natural killer cell activation inhibition; organ transplantation.

XX Homo sapiens.

OS Synthetic.

PN US6162434-A.

PD 19-DEC-2000.

PE 03-MAY-1995; 95US-0433613.

PR 03-MAY-1995; 95US-0433613.

PA (SANG-) SANGSTAT MEDICAL CORP.

XX Buelow R;

DR WPI; 2001-111720/12.

XX Inhibiting activation of cytotoxic T lymphocytes and/or natural killer

PT cells involves combining cells with compound comprising oligopeptide

PT with contiguous sequence of human leukocyte antigen-B alpha1 domain

PS Disclosure; Column 7-8; 20pp; English.

XX The present sequence is a peptide which may be used in a method for

CC inhibiting activation of cytotoxic T lymphocytes and/or natural killer

CC cells. The method involves combining the cells with a compound

CC comprising an oligopeptide of at least 6 amino acids comprising a

CC contiguous sequence of a human leukocyte antigen (HLA)-B alpha1 domain

CC including amino acids 84-86, where amino acids 84-86 are YVW.

CC The method is useful for transplanting a donor mammalian organ or cells

CC to a mammalian recipient. The organ or cells are combined with the

CC compound prior to implanting in the mammalian recipient, or the compound

CC may be administered to the mammalian recipient in a period extending

DB 7 RIALRY 13

RESULT 7
 AAB59418
 ID AAB59418 standard; Peptide; 14 AA.

XX AAB59418;

DT 22-MAR-2001 (first entry)

DE Human Class I HLA-B alpha1-domain-derived peptide #18.

XX Human; human leukocyte antigen-B; HLA-B; immunosuppressive;

KW cytotoxic T cell activation inhibition; cell transplantation;

KW natural killer cell activation inhibition; organ transplantation.

XX Homo sapiens.

OS Synthetic.

PN US6162434-A.

PD 19-DEC-2000.

PE 03-MAY-1995; 95US-0433613.

PR 03-MAY-1995; 95US-0433613.

PA (SANG-) SANGSTAT MEDICAL CORP.

XX Buelow R;

DR WPI; 2001-111720/12.

XX Inhibiting activation of cytotoxic T lymphocytes and/or natural killer

PT cells involves combining cells with compound comprising oligopeptide

PT with contiguous sequence of human leukocyte antigen-B alpha1 domain

PS Disclosure; Column 7-8; 20pp; English.

XX The present sequence is a peptide which may be used in a method for

CC inhibiting activation of cytotoxic T lymphocytes and/or natural killer

CC cells. The method involves combining the cells with a compound

CC comprising an oligopeptide of at least 6 amino acids comprising a

CC contiguous sequence of a human leukocyte antigen (HLA)-B alpha1 domain

CC including amino acids 84-86, where amino acids 84-86 are YVW.

CC The method is useful for transplanting a donor mammalian organ or cells

CC to a mammalian recipient. The organ or cells are combined with the

CC compound prior to implanting in the mammalian recipient, or the compound

CC may be administered to the mammalian recipient in a period extending

CC from prior to implanting the donor organ or cells.

SO Sequence 14 AA;

Query Match 62.1%; Score 36; DB 22; Length 14;
 Best Local Similarity 100.0%; Pred. No. 3.1;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RIALRY 7
 DB 6 RIALRY 12

RESULT 8
 AAR92911

ID AAR92911 standard; peptide; 20 AA.

XX AAR92911;

DT 16-MAY-1996 (first entry)

DE HLA-B2702 CTL modulating peptide (B2702.84-75/84-75).

XX Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;
 KW immunosuppressant; graft versus host disorder; transplantation; therapy;
 KW Class I MHC; HLA-B2702.
 XX Synthetic.
 OS
 PN WO9526979-A1.
 XX
 PD 12-OCT-1995.
 XX
 XX PF 05-APR-1995; 95WO-US04349.
 XX
 XX PR 05-APR-1994; 94US-0222851.
 XX
 XX PA (STRD) UNIV LELAND STANFORD JUNIOR.
 XX
 XX PI Clayberger C, Krensky AM, Parham P;
 XX WPI; 1995-358582/46.
 XX
 XX Extension of acceptance period of transplants from MHC unmatched
 PT donor hosts - using Class I B75-84 MHC antigen of the recipient
 PT host
 XX
 XX Example 15; Page 36; 80pp; English.
 XX
 CC AAR33061-R83085, AAR33090-R83096 and AAR32907-R32914 represent fragments
 CC of class I major histocompatibility complex (MHC) antigens. This
 CC sequence is a dimer of residues 75-84 of the alpha-1 domain of the class
 CC I MHC HLA-B2702. These sequences can be used to extend the period of
 CC acceptance by a recipient of a transplant from an MHC unmatched donor.
 CC The peptides are administered to a patient in conjunction with a
 CC subtherapeutic amount of an immunosuppressant. This is administered to
 CC the patient for a limited period of time (compared to the lifetime
 CC administration for current treatments). The peptides particularly
 CC modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)
 CC of the patient.
 XX
 SQ Sequence 20 AA;
 Query Match 59.5%; Score 34.5; DB 16; Length 20;
 Best Local Similarity 60.0%; Pred. No. 8.7;
 Matches 9; Conservative 2; Mismatches 1; Indels 3; Gaps 1;
 QY 1 RIALRY---YRLAIR 12
 Db |.:|:| |
 2 RLAILRLNRYRLAIR 16
 RESULT 9
 AAW33779
 ID AAW33779 standard; peptide; 20 AA.
 XX
 AC AAW33779;
 XX
 DT 19-JUN-1998 (first entry)
 XX
 DE Immunomodulating dimer peptide #2.
 XX
 KW Immunomodulating dimer; immunosuppressant drug; CTL activation;
 KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;
 KW rejection.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 PN WO9744351-A1.
 XX
 XX PD 27-NOV-1997.
 XX
 XX PF 22-MAY-1997; 97WO-US08689.
 XX

PR 24-MAY-1996; 96US-0653294.
 XX
 XX PA (STRD) UNIV LELAND STANFORD JUNIOR.
 XX
 XX PI Beulow R, Clayberger C, Krensky AM;
 XX WPI; 1998-086530/08.
 XX
 XX New immunomodulating dimer peptide(s) - based on a Class I HLA-B
 PT alpha-1 domain, used for preventing rejection of transplants or
 PT treating autoimmune diseases
 XX
 XX Claim 16; Page 35; 41pp; English.
 XX
 CC This sequence represents a specifically claimed immunomodulating
 CC dimer peptide of the invention. A peptide-type compound or variant is
 CC claimed which has immunomodulating activity, including the N-terminal
 CC acylated and/or C-terminal amidated or esterified forms of up to 60
 CC amino acids, where the peptide-type compound comprises the formula; A-B,
 CC where A, B = (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or
 CC V; aa77 = D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a
 CC hydrophobic or small amino acid; aa82 = R or L; aa83 = G or R; and aa
 CC represents amino acid. The sequence in the brackets may optionally be
 CC absent or truncated at any peptide type bond within the brackets. The
 CC compounds comprise amino acid sequences related to a Class I HLA-B
 CC alpha domain (positions 79-84). They can be used to inhibit cytotoxic
 CC T-lymphocytes (CTL) from undesirably attacking cells in a host or in
 CC vitro. They can also be used in combination with antigenic peptides or
 CC proteins of interest to activate CTLs. They can also inhibit the
 CC proliferation of T cells in response to anti-CD3. The peptide can be
 CC used for preventing rejection of transplants or for treating autoimmune
 CC diseases, e.g. diabetes, rheumatoid arthritis and lupus erythematosus.
 CC The products can also be used for detection and diagnosis.
 XX
 SQ Sequence 20 AA;
 Query Match 59.5%; Score 34.5; DB 19; Length 20;
 Best Local Similarity 60.0%; Pred. No. 8.7;
 Matches 9; Conservative 2; Mismatches 1; Indels 3; Gaps 1;
 QY 1 RIALRY---YRLAIR 12
 Db |.:|:| |
 2 RLAILRLNRYRLAIR 16
 RESULT 10
 AAW83083
 ID AAW83083 standard; Protein; 46 AA.
 XX
 AC AAW83083;
 XX
 DT 07-NOV-2001 (first entry)
 XX
 DE Human immune/haematopoietic antigen SEQ ID NO:10676.
 XX
 KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
 KW cytostatic; gene therapy; vaccine; metastasis.
 XX
 OS Homo sapiens.
 XX
 PN WO200157182-A2.
 XX
 XX PD 09-AUG-2001.
 XX
 XX PF 17-JAN-2001; 2001WO-US01354.
 XX
 XX PR 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.

```
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 14-AUG-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 06-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0233397.
PR 14-SEP-2000; 2000US-0233398.
PR 14-SEP-2000; 2000US-0233399.
PR 14-SEP-2000; 2000US-0233400.
PR 14-SEP-2000; 2000US-023401.
PR 14-SEP-2000; 2000US-0235063.
PR 14-SEP-2000; 2000US-0235064.
PR 14-SEP-2000; 2000US-0235065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.

PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 01-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.
XX
PA
PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001.483426/52.
XX
XX N-PSDB; AAK55864.
XX
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
XX useful for preventing, diagnosing and/or treating cancers and
XX metacasts -
XX
XX
XX Claim 11; SEQ ID NO 10676; 3071pp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
XX amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
XX activity, and can be used in gene therapy and vaccine production. (I)
XX
```

Sequences AAY87551-Y87564 and AAY87568 represent novel analogues of peptide YY (PYY) that can be used for the treatment of gastrointestinal disorders. PYY (AAY87549, AAY87550) is a 36 residue peptide amide that is released into the circulation after a meal and which is thought to play a role in regulating intestinal secretion and absorption. It binds

DR N-PSDB; AAS98485.

XX

PT Polypeptides and polynucleotides comprising protective sequences useful for preventing, delaying or rescuing a cell from death in disease, condition or disorders such as Alzheimer's disease, stroke, tumours, PT

PT

```
PT trauma -
XX Claim 17; Fig 6AU; 228bp; English.
PS
XX The invention relates to an isolated polypeptide encoded by a protective
XX sequence, which is a polynucleotide comprising sequences which, when
XX introduced into a cell either predisposed to undergo cell death or in the
XX process of undergoing cell death, prevent, delay, or rescue the cell from
XX death, relative to a corresponding cell into which exogenous nucleic
XX acids have been introduced. The sequences of the invention are useful for
XX diagnosing a protective sequence-mediated condition, disorder or disease
XX in an individual. The treatable disorders are preferably of the central
XX nervous system of humans including ischaemia-related conditions such as
XX stroke, cerebral herniation, septic embolism, cerebral oedema, infections
XX such as meningitis, protozoal infections such as malaria, metazoal
XX infections such as echinococcosis, vascular diseases such as ischaemic
XX encephalopathy, conditions involving the eye such as macular
XX degeneration, diabetic retinopathy, trauma such as epidural haematoma,
XX tumours such as primary intracranial tumours, degenerative diseases such
XX as Alzheimer's disease and nutritional, environmental and metabolic
XX conditions. Sequences AAU73255-AAU73378 represent open reading frames of
XX the human protective sequence polynucleotides.
SQ Sequence 36 AA:

Query Match 55.2%; Score 32; DB 23; Length 36;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 LRYRL 9
Db 19 LRYRL 24

RESULT 13
AAU53267
ID AAU53267 standard; Protein; 57 AA.
XX
XX AAU53267;
AC
XX
XX 27-FEB-2002 (first entry)
DT
XX
XX Propionibacterium acnes immunogenic protein #14163.
DE
XX
XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
XX uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
XX inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
XX dermatological; osteopathic; neuroprotectant.
XX
XX Propionibacterium acnes.
OS
XX
XX WO200181581-A2.
PN
XX
XX 01-NOV-2001.
PD
XX
XX 20-APR-2001; 2001WO-US12865.
PF
XX
XX 21-APR-2000; 2000US-199047P.
PR
XX 02-JUN-2000; 2000US-208841P.
PR 07-JUL-2000; 2000US-216747P.
PR
XX
XX (CORI-) CORIXA CORP.
PA
XX
XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
XX L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
XX WPI; 2001-616774/71.
DR
XX
XX N-PSDB; AAS59559.
XX
XX Propionibacterium acnes polypeptides and nucleic acids useful for
XX PT vaccinating against and diagnosing infections, especially useful for
XX PT treating acne vulgaris -
XX
```

```
PS Example 1; SEQ ID No 14462; 1069bp; English.
XX
XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
XX polypeptides. The proteins and their associated DNA sequences are used in
XX the treatment, prevention and diagnosis of medical conditions caused by
XX P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
XX pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
XX P. acnes is also involved in infections of bone, joints and the central
XX nervous system, however it is particularly involved in the inflammatory
XX lesions associated with acne vulgaris. A method for detecting the
XX presence or absence of P. acnes in a patient comprises contacting a
XX sample with a binding agent that binds to the proteins of the invention
XX and determining the amount of bound protein in the sample. The
XX polypeptides may be used as antigens in the production of antibodies
XX specific for P. acnes proteins. These antibodies can be used to
XX downregulate expression and activity of P. acnes polypeptides and
XX therefore treat P. acnes infections. The antibodies may also be used as
XX diagnostic agents for determining P. acnes presence, for example, by
XX enzyme linked immunosorbent assay (ELISA).
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
SQ Sequence 57 AA:

Query Match 55.2%; Score 32; DB 22; Length 57;
Best Local Similarity 62.5%; Pred. No. 79;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 5 RYRLAIR 12
Db 1 RYRLAIR 8

RESULT 14
AAB59419
ID AAB59419 standard; Peptide; 13 AA.
XX
XX AAB59419;
AC
XX
XX 22-MAR-2001 (first entry)
DT
XX
XX Human Class I HLA-B alpha1-domain-derived peptide #19.
DE
XX
XX Human; human leukocyte antigen-B; HLA-B; immunosuppressive;
XX cytotoxic T cell activation inhibition; cell transplantation;
XX natural killer cell activation inhibition; organ transplantation.
XX
XX Homo sapiens.
OS
XX
XX Synthetic.
PN
XX
XX US6162434-A.
PN
XX
XX 19-DEC-2000.
PD
XX
XX 03-MAY-1995; 95US-0433613.
PF
XX
XX 03-MAY-1995; 95US-0433613.
PR
XX
XX (SANG-) SANGSTAT MEDICAL CORP.
PA
XX
XX Buelow R;
XX
XX WPI; 2001-111720/12.
DR
XX
XX Inhibiting activation of cytotoxic T lymphocytes and/or natural killer
XX PT cells involves combining cells with compound comprising oligopeptide
XX PT with contiguous sequence of human leukocyte antigen-B alpha1 domain -
XX
XX Disclosure; Column 7-8; 20pp; English.
XX
XX The present sequence is a peptide which may be used in a method for
XX inhibiting activation of cytotoxic T lymphocytes and/or natural killer
XX
```

CC cells. The method involves combining the cells with a compound
 CC comprising an oligopeptide of at least 6 amino acids comprising a
 CC contiguous sequence of a human leukocyte antigen (HLA)-B alpha1 domain
 CC including amino acids 84-86, where amino acids 84-86 are YYW.
 CC The method is useful for transplanting a donor mammalian organ or cells
 CC to a mammalian recipient. The organ or cells are combined with the
 CC compound prior to implanting in the mammalian recipient, or the compound
 CC may be administered to the mammalian recipient in a period extending
 CC from prior to implanting the donor organ or cells.

XX SQ Sequence 13 AA;

Query Match 53.4%; Score 31; DB 22; Length 13;

Best Local Similarity 100.0%; Pred. No. 24;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 IALRY 7

|||||

Db 5 IALRY 10

RESULT 15

AAW99406

ID AAW99406 standard; peptide; 18 AA.

AC AAW99406;

DT 08-JUN-1999 (first entry)

XX Protegrin derivative peptide SM2187.

DE Linear; protegrin; peptide antibiotic; beta-sheet; secondary structure;

KW disulphide bridge; antibody; ribozyme; antitumour agent; antiviral;

KW anti-inflammatory; mammal; cell membrane; passive transport; cytoplasm;

KW nucleus; blood-brain barrier.

XX Synthetic.

OS WO9907728-A2.

PN 18-FEB-1999.

XX 06-AUG-1998; 98WO-FR01757.

PF 12-AUG-1997; 97FR-0010297.

XX (SYNT-) SYNT:EM SA.

PA Calas B, Chavanieu A, Grassy G, Kaczorek M;

PI WPI; 1999-190034/16.

XX Derivatives of antibiotic peptides lacking disulfide bridges - used

PT as carriers to deliver active agents into cells

XX Claim 7; Page 28; 37pp; French.

PS This peptide represents a linear derivative of the protegrin family of

XX peptide antibiotics. Protegrin antibiotics form part of the peptide

CC antibiotic family which contain a beta-sheet secondary structure linked

CC by disulphide bridges. The new derivatives are linear and lack the

CC disulphide bridge. The novel derivatives are used to deliver active

CC agents to an organism, e.g. therapeutic proteins, antibodies (or their

CC fragments), nucleic acid, oligonucleotides, ribozymes, antitumour agents,

CC antivirals and anti-inflammatories, etc. The derivatives are non-toxic

CC and non-lytic but can cross mammalian cell membranes rapidly by a passive

CC mechanism, so can deliver active agents to cytoplasm and nucleus,

CC including crossing the blood-brain barrier.

XX SQ Sequence 18 AA;

Query Match 53.4%; Score 31; DB 20; Length 18;

Best Local Similarity 58.3%; Pred. No. 35;

Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 RIALRYRLAIR 12

||:| | |||

Db 4 RLAVRLRLFAIR 15

Search completed: December 19, 2002, 16:19:07

Job time : 29.5882 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 19, 2002, 16:14:32 ; Search time 21.1765 Seconds
(without alignments)
116.760 Million cell updates/sec

Title: US-08-653-294C-40
Perfect score: 58
Sequence: 1 RIALRYRLAIR 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 45785

Minimum DB seq length: 0
Maximum DB seq length: 60

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPREMBL 21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_muc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	31	53.4	33	16	O8XFT3 salmonella
2	31	53.4	56	17	O27979 archaeoglob
3	30	51.7	31	6	Q28493 mirounga le
4	30	51.7	47	5	Q9U9S2 trichinella
5	30	51.7	52	16	Q92HL9 rickettsia
6	30	51.7	53	4	Q9P138 homo sapien
7	29	50.0	35	8	Q951R3 anthromurice
8	29	50.0	53	17	O8ZV69 pyrobaculum
9	29	50.0	56	5	Q8T946 drosophila
10	29	50.0	60	16	Q9ZJP5 rickettsia
11	28	48.3	25	10	Q9S8T8 coloccasia e
12	28	48.3	35	12	Q65737 bluetongue
13	28	48.3	40	16	Q9PC59 xylella fas
14	28	48.3	57	5	Q9NLE7 leishmania
15	28	48.3	57	12	Q8V7E7 tt virus. o
16	27	46.6	23	8	Q32615 marchantia

17	27	46.6	27	4	O95596	homo sapien
18	27	46.6	42	3	O95167	saccharomyc
19	27	46.6	44	5	Q9TWM4	lymaaea sta
20	27	46.6	45	16	P75718	escherichia
21	27	46.6	52	5	O44363	ceratitis c
22	27	46.6	58	16	Q98E76	rhizobium l
23	27	46.6	60	2	Q50087	mycobacteri
24	26	44.8	19	12	O90630	baboon herp
25	26	44.8	19	12	O90633	baboon herp
26	26	44.8	30	7	Q95HE9	papio anubi
27	26	44.8	30	7	Q95HE8	papio anubi
28	26	44.8	30	7	Q95HE7	papio anubi
29	26	44.8	30	7	Q95HE6	papio anubi
30	26	44.8	30	7	Q95HE5	papio anubi
31	26	44.8	30	7	Q95HE4	papio anubi
32	26	44.8	30	7	Q95HE3	papio anubi
33	26	44.8	30	7	Q95HE1	papio anubi
34	26	44.8	30	7	Q95HE0	papio anubi
35	26	44.8	30	7	Q95H96	papio anubi
36	26	44.8	30	7	Q95H95	papio anubi
37	26	44.8	30	7	Q95H94	papio anubi
38	26	44.8	39	7	O19688	homo sapien
39	26	44.8	40	4	P78340	homo sapien
40	26	44.8	43	10	Q9S8M4	tritricum ae
41	26	44.8	45	10	O49773	arabidopsis
42	26	44.8	52	10	Q9SK54	arabidopsis
43	26	44.8	53	16	Q92ZA6	rhizobium m
44	26	44.8	54	12	Q9W8Y9	simian viru
45	26	44.8	57	5	Q9W132	drosophila

ALIGNMENTS

RESULT 1

O8XFT3 PRELIMINARY; PRT; 33 AA.
ID Q8XFT3
AC O8XFT3;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Putative cytoplasmic protein (Hypothetical protein STY0391).
GN STY0359 OR STY0391.
OS Salmonella typhimurium, and
OS Salmonella typhi.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella
OX NCBI_TaxID=602, 601;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhimurium; STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Letreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2.";
RL Nature 413:852-856(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhi; STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Bentley S.D., Holden M.T.G., Sebaihia M.,
RA Churcher C., Mungall K.L., Brooks K., Chillingworth T., Connor P.,
RA Baker S., Basham D., Brooks K., Davies R.M., White N., Farrar J.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrrell B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella

```

RN Enterica serovar Typhi CT18."
RL Nature 413:848-852(2001).
DR EMBL; AE008711; AAL19313.1; -.
DR EMBL; AL627266; CAD0814.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 33 AA; 3814 MW; 8242360061FCED1 CRC64;

Query Match
Best Local Similarity 53.4%; Score 31; DB 16; Length 33;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 5 RYRYRLAI 11
DB 4 RYRYRAV 10

RESULT 2
O27979 PRELIMINARY; PRT; 56 AA.
ID O27979
AC O27979;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Hypothetical protein AF2305.
GN AF2305.
OS Archaeoglobus fulgidus.
OC Archaeae; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC Archaeoglobaceae; Archaeoglobus.
OX NCB1_TaxID=22234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kierlavage A.R., Graham D.E., Kyriades N.C.,
RA Fleischman R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kitzness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Uteerback T.,
RA Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
reducing archaeon Archaeoglobus fulgidus."
RL Nature 390:364-370(1997).
DR EMBL; AE000945; AAB88955.1; -.
DR TIGR; AF2305; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 56 AA; 6986 MW; DBD220C268A7DF6 CRC64;

Query Match
Best Local Similarity 53.4%; Score 31; DB 17; Length 56;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 3 ALRYRLAIR 12
DB 32 AERYRLAIR 41

RESULT 3
O28493 PRELIMINARY; PRT; 31 AA.
ID O28493
AC O28493;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-UN-2002 (TREMBLrel. 21, Last annotation update)
DE Histone H2AF (Fragment).
OS Mironuga leonina (Southern elephant seal).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euteria; Carnivora; Plinipedia; Phocidae; Mironuga.
OX NCB1_TaxID=9715;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94214718; PubMed=7909260;
RA Slide R.W., Moritz C., Heideman A., Hale P.T.;
RT "Rapid assessment of single-copy nuclear DNA variation in diverse
species."
RL Mol. Ecol. 2:359-373(1993).
DR EMBL; L17501; AAA72378.1; -.
DR InterPro; IPR004822; Histone_core.
DR Pfam; PF00125; histone; 1.
FT NON TER 1
FT NON TER 1
SQ SEQUENCE 31 AA; 3400 MW; 01C0E585ABF3259E CRC64;

Query Match
Best Local Similarity 51.7%; Score 30; DB 6; Length 31;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 RLALRYRLAIR 12
DB 4 RITPRHLQLAIR 15

RESULT 4
O90982 PRELIMINARY; PRT; 47 AA.
ID O90982
AC O90982;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE Hypothetical 5.6 kDa protein (Fragment).
OS Trichinella pseudospiralis.
OC Eukaryota; Metazoa; Nematoda; Enoplea; Trichocephalida;
OC Trichinellidae; Trichinella.
OX NCB1_TaxID=6337;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TPR(L);
RA Sengenova S.K., Chrisanova G.G.;
RT "RAPD-derived Trichinella pseudospiralis probe."
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF153496; AAD45671.1; -.
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 47 AA; 5588 MW; 7483E6A1213062 CRC64;

Query Match
Best Local Similarity 51.7%; Score 30; DB 5; Length 47;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 2 IALRYRLAIR 12
DB 25 VASRYRLAIRIQ 35

RESULT 5
O92HL9 PRELIMINARY; PRT; 52 AA.
ID O92HL9
AC O92HL9;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Hypothetical protein RC0752.
GN RC0752.
OS Rickettsia conorii.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Rickettsiae; Rickettsia.
OX NCB1_TaxID=781;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MALISH 7;
RX MEDLINE=21442074; PubMed=11557893;
RA Ogata H., Audic S., Renesto-Audiffren P., Fournier P.-E., Barbe V.,

```


RA Samson D., Roux V., Cossart P., Weissenbach J., Claverie J.-M.,
 RA Racult D.;
 RT "Mechanisms of evolution in Rickettsia conorii and R. prowazekii.";
 RL Science 293:2093-2098(2001).
 DR EMBL: AE008632; AAL03290.1; -.
 DR InterPro: IPR001440; TPR.
 SQ SEQUENCE 52 AA; 6293 MW; 5A3C8F20C939BD36 CRC64;
 Query Match 51.7%; Score 30; DB 16; Length 52;
 Best Local Similarity 50.0%; Pred. No. 2.1e+02;
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 IALRYRLAI 11
 :|||:
 DB 11 LAIKYNSAI 20

RESULT 6
 Q9P138 PRELIMINARY; PRT; 53 AA.
 ID Q9P138;
 AC Q9P138;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE PRO2964.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RA Zhang C., Yu Y., Zhang S., Wei H., Zhang Y., Zhou G., Bi J., Liu M.,
 RA He F.;
 RT "Functional prediction of the coding sequences of 79 new genes deduced
 RT by analysis of cDNA clones from human fetal liver."
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF119910; AAF69664.1; -.
 SQ SEQUENCE 53 AA; 6734 MW; 512FCBB61EA08BC CRC64;

Query Match 51.7%; Score 30; DB 4; Length 53;
 Best Local Similarity 62.5%; Pred. No. 2.2e+02;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 IALRYRL 9
 :|||:
 DB 35 VANRYRI 42

RESULT 7
 Q951R3 PRELIMINARY; PRT; 35 AA.
 ID Q951R3;
 AC Q951R3;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Muts-like protein (Fragment).
 GN MSH.
 OS Anthomuricea sp. USNM-94442.
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Alcyonaria; Gorgonacea;
 OC Holaxonia; Plexauridae; Anthomuricea.
 OX NCBI_TaxID=161894;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CR105-5;
 RA France S.C., Hoover L.L.;
 RT "Analysis of variation in mitochondrial DNA sequences (ND3, ND4L, MSH)
 RT among Octocorallia (Cnidaria: Anthozoa).";
 RL Proc. Biol. Soc. Wash. 0:0-0(2001).
 DR EMBL: AF385304; AAK68137.1; -.
 KW Mitochondrion.

FT NON TER 35 35
 SQ SEQUENCE 35 AA; 4214 MW; F41BC3BDE6F9B215 CRC64;
 Query Match 50.0%; Score 29; DB 8; Length 35;
 Best Local Similarity 50.0%; Pred. No. 2.2e+02;
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 RIALRYRLA 10
 :|||:
 DB 3 QIPMYFNLA 12

RESULT 8
 Q8ZV69 PRELIMINARY; PRT; 53 AA.
 ID Q8ZV69;
 AC Q8ZV69;
 DT 01-MAR-2002 (TREMBlrel. 20, Created)
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Hypothetical protein PAE2427.
 GN PAE2427.
 OS Pyrobaculum aerophilum.
 OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
 OC Thermoproteaceae; Pyrobaculum.
 OX NCBI_TaxID=13773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=IM2 / ATCC 51768 / DSM 7523;
 RX PubMed=11792869;
 RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
 RA Miller J.H.;
 RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
 RT aerophilum";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).
 DR EMBL: AE009875; AAL64187.1; -.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 53 AA; 6120 MW; 73023B57A6A8E7F CRC64;

Query Match 50.0%; Score 29; DB 17; Length 53;
 Best Local Similarity 54.5%; Pred. No. 3.3e+02;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 RIALRYRLAI 11
 :|||:
 DB 35 RIAMWYLQAL 45

RESULT 9
 Q8T946 PRELIMINARY; PRT; 56 AA.
 ID Q8T946;
 AC Q8T946;
 DT 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE GM10545p.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Friese E.,
 RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
 RA Miranda A., Mungall C.J., Nunoo J., Paclab J., Paragas V., Park S.,
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
 RA Celniker S.;
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY070857; AAL48479.1; -.
 SQ SEQUENCE 56 AA; 6512 MW; E9803230DB63B169 CRC64;

Query Match 50.0%; Score 29; DB 5; Length 56;

Best Local Similarity 55.6%; Pred. No. 3.5e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 RIALRYRL 9
| | | | |
7 RSALRPFR 15

RESULT 10

Q92JP5 PRELIMINARY; PRT; 60 AA.

AC Q92JP5; 01-DEC-2001 (TRENBLrel. 19, Created)

DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)

DE 01-JUN-2002 (TRENBLrel. 21, Last annotation update)

GN RCO022. Hypothetical protein RCO022.

OS Rickettsia conorii.

OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;

OC Rickettsiaceae; Rickettsiae; Rickettsia.

OX NCBI_TaxID=781;

RP SEQUENCE FROM N.A.

RC STRAIN=MALISH 7;

RX MEDLINE=21442074; PubMed=11557893;

RA Ogata H., Audic S., Renesto-Audifren P., Fournier P.-E., Barbe V.,

RA Samsom D., Roux V., Cossart P., Weissenbach J., Claverie J.-M.,

RA Raoult D.;

RT "Mechanisms of evolution in Rickettsia conorii and R. prowazekii."

RL Science 293:2093-2098 (2001).

DR EMBL, AF008571; AA02560.1; -.

DR InterPro; IPR001440; TPR.

DR SMART; SM00028; TPR; 1.

KW Hypothetical protein; Complete proteome.

SV SEQUENCE 60 AA; 7160 MW; 4B5F2B0B5FB7021C CRC64;

Query Match 50.0%; Score 29; DB 16; Length 60;
Best Local Similarity 40.0%; Pred. No. 3.8e+02;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 3 ALRYRLAIR 12
| | | | |
27 SLKYKXKX 36

RESULT 11

Q9S8T8 PRELIMINARY; PRT; 25 AA.

AC Q9S8T8; 01-MAY-2000 (TRENBLrel. 13, Created)

DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)

DE 01-OCT-2001 (TRENBLrel. 18, Last annotation update)

DE Globulin G2B (Fragment).

OS Colocasia esculenta (Blighant's ear) (Taro).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Araceae; Colocasia.

OX NCBI_TaxID=4460;

RP SEQUENCE.

RX MEDLINE=93104679; PubMed=1467653;

RA de Castro L.A., Carneiro M., Neshion D de C., de Paiva G.R.,

RT "Spectral and temporal gene expression patterns occur during corn

RL plant Cell 4:1549-1559 (1992).

DR InterPro; IPR002160; Kunitz legume.

DR Pfam; PF00197; Kunitz legume; 1.

SV SEQUENCE 25 AA; 2861 MW; 409579EB835C793A CRC64;

Query Match 48.3%; Score 28; DB 10; Length 25;
Best Local Similarity 50.0%; Pred. No. 2.4e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 5 RYRLAIR 12

DB 17 RYRLAIR 24
| | | | |

RESULT 12

Q65737 PRELIMINARY; PRT; 35 AA.

AC Q65737; 01-NOV-1996 (TRENBLrel. 01, Created)

DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)

DE 01-NOV-1998 (TRENBLrel. 08, Last annotation update)

DE Serotype-specific antigen (S2) RNA, 5' end (Fragment).

OS Bluetongue virus.

OC Viruses; dsRNA viruses; Reoviridae; Orbivirus.

OX NCBI_TaxID=12591;

RP SEQUENCE FROM N.A.

RX MEDLINE=88072081; PubMed=2825414;

RA Huisman H., Cloete M., le Roux A.;

RT "The genetic relatedness of a number of individual cognate genes of

RL viruses in the bluetongue and closely related serogroups."

RL Virology 161:421-428 (1987).

DR EMBL; M22442; AAA42841.1; -.

FT NON TER 35 35

SV SEQUENCE 35 AA; 4186 MW; C7E02BAA0510DF5 CRC64;

Query Match 48.3%; Score 28; DB 12; Length 35;
Best Local Similarity 66.7%; Pred. No. 3.4e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 LRYRLAIR 12
| | | | |
19 LNHPLAIR 27

RESULT 13

Q9PC59 PRELIMINARY; PRT; 40 AA.

AC Q9PC59; 01-OCT-2000 (TRENBLrel. 15, Created)

DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)

DE 01-MAR-2002 (TRENBLrel. 20, Last annotation update)

DE Hypothetical protein Xf1922.

OS Xylella fastidiosa.

OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;

OC Xylella.

OX NCBI_TaxID=2371;

RP SEQUENCE FROM N.A.

RC STRAIN=9A5C;

RX MEDLINE=20365717; PubMed=10910347;

RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,

RA Alvaranga R., Alves L.M.C., Araya J.E., Bata G.S., Baptista C.S.,

RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Brites M.R.S.,

RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,

RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,

RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,

RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,

RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,

RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,

RA Ho P.L., Hohnsels J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,

RA Krieger J.E., Kuramae B.E., Laigret F., Lambais M.R., Leite L.C.C.,

RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,

RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,

RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,

RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,

RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Neto L.E.S.,

RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,

RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,

RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,

RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,

RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,

RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Teai S.M., Tsuchiko M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT "The genome sequence of the plant pathogen *Xylella fastidiosa*.";
RL Nature 406:151-159(2000)
DR EMBL; AB004012; AAF84728.1; -;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 40 AA; 4533 MW; 96296103833C1B08 CRC64;

Query Match 48.3%; Score 28; DB 16; Length 40;
Best Local Similarity 75.0%; Pred. No. 3.8e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 ALRYRLA 10
| | | | |
Db 30 AARYYDLA 37

RESULT 14
Q9NLE7
ID Q9NLE7 PRELIMINARY; PRT; 57 AA.
AC Q9NLE7;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE Hypothetical 6.5 kDa protein (Fragment).
GN LM26.495.
OS Leishmania major.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FRIEDLIN;
RA Murphy L., Quail M., Harris D., Rajandream M., Ivens A., Barrell B.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL160493; CAB98113.1; -;
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 57 AA; 6498 MW; 6DF143AEF1C8C0D1 CRC64;

Query Match 48.3%; Score 28; DB 5; Length 57;
Best Local Similarity 55.6%; Pred. No. 5.4e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 LRYRLAIR 12
| | | | |
Db 37 LVYRLIVK 45

RESULT 15
Q8VTE7
ID Q8VTE7 PRELIMINARY; PRT; 57 AA.
AC Q8VTE7;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE ORF1 (Fragment).
OS TT virus.
OC Viruses; ssDNA viruses; unclassified ssDNA viruses.
OX NCBI_TaxID=68887;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=2184401; PubMed=11855633;
RA Peng Y.H., Nishizawa T., Takahashi M., Ishikawa T., Yoshikawa A.,
RA Okamoto H.;
RT "Analysis of the complete genomes of thirteen TT virus variants
RT classifiable into the fourth and fifth genetic groups, isolated from
RT viremic infants.";
RL Arch. Virol. 147:21-41(2002).
DR EMBL; AB064612; BAB79368.1; -;
FT NON_TER 57

SQ SEQUENCE 57 AA; 8138 MW; 61E467DDE395ACE9 CRC64;
Query Match 48.3%; Score 28; DB 12; Length 57;
Best Local Similarity 62.5%; Pred. No. 5.4e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 5 RYRLAIR 12
| | | | |
Db 48 RYRLAIR 55

Search completed: December 19, 2002, 16:21:53
Job time : 22.1765 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 19, 2002, 16:03:12 ; Search time 5.47059 Seconds
(without alignments)
90.980 Million cell updates/sec

Title: US-08-653-294C-40
Perfect score: 58
Sequence: 1 RIALRYRLAIR 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues 5116
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 60

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	63.8	44	1 YIM8 BPPH1	P10431 bacteriophage
2	28	48.3	24	1 LPA2 BACSU	P08497 bacillus su
3	27	46.6	30	1 PLMS SQUAC	P82542 squalus aca
4	26	44.8	31	1 DEF2 MESAU	P81466 mesocricetu
5	26	44.8	33	1 DBF4 MESAU	P81468 mesocricetu
6	26	44.8	42	1 Y281 TREPA	O83305 treponema p
7	25	43.1	56	1 LCRT YERPS	Q00932 yersinia ps
8	25	43.1	60	1 RS21 MYCPN	P57079 mycoplasma
9	24	41.4	36	1 PSBY PORPU	P51206 porphyra pu
10	24	41.4	50	1 HSP1 HUMAN	P04553 homo sapien
11	24	41.4	50	1 HSP1 PONYF	P35310 pongo pygma
12	24	41.4	51	1 Y331 BORBU	O51309 borrelia bu
13	24	41.4	55	1 Y072 NPVOP	O10325 oryza pseu
14	23	39.7	26	1 MGN CHICK	P50594 gallus gall
15	23	39.7	31	1 PSAM CYAPA	P48185 cyanophora
16	23	39.7	36	1 PMY PETWA	P80024 petromyzon
17	23	39.7	45	1 Y00D BPT4	Q01434 bacterioph
18	23	39.7	57	1 IVBK DENAN	P00982 dendroaspis
19	23	39.7	58	1 LHAL RHOSH	P02949 rhodobacter
20	23	39.7	59	1 RL32 BORBU	O51646 borrelia bu
21	22	37.9	21	1 HPAA STRGD	P80437 streptomyce
22	22	37.9	26	1 AMPM BONMO	P81495 bombyx mori
23	22	37.9	29	1 PSAM GUITH	O78448 guillardia
24	22	37.9	36	1 NEUY GADMO	P80167 gadus morhu
25	22	37.9	36	1 NEUY ONCMY	P29071 oncorhynch
26	22	37.9	36	1 NEUY FIG	P01304 sus scrofa
27	22	37.9	36	1 NEUY RABIT	P09640 oryctolagus
28	22	37.9	36	1 NEUY RANRI	P29949 rana ridibu
29	22	37.9	36	1 NEUY SHEEP	P14765 ovis aries
30	22	37.9	36	1 PYI AMICA	P29205 amia calva
31	22	37.9	36	1 PYI LEPSP	P09473 lepisosteus
32	22	37.9	36	1 PYI ONCKI	P09474 oncorhynch
33	22	37.9	36	1 PYI RAJRH	P29206 raja rhina

ALIGNMENTS

RESULT 1
YIM8_BPPH1
ID YIM8_BPPH1 STANDARD; PRT; 44 AA.
AC P10431;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 01-MAR-1989 (Rel. 10, Last annotation update)
DE Hypothetical immunity region protein 8.
OS Bacteriophage phi-105.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
OC Lambda-like viruses.
OX NCBI_TaxID=10717;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86056972; PubMed=3934047;
RA Cully D.F., Garro A.J.
RT "Nucleotide sequence of the immunity region of Bacillus subtilis
bacteriophage phi 105: identification of the repressor gene and its
mRNA and protein products.";
RL Gene 38:153-164(1985).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M11920; ; NOT_ANNOTATED_CDS.
DR F01; G24521; IMBP8.
KW Hypothetical protein.
SQ SEQUENCE 44 AA; 5574 MW; C440D1B134841832 CRC64;
Query Match 63.8%; Score 37; DB 1; Length 44;
Best Local Similarity 54.5%; Pred. No. 0.54; Indels 0; Gaps 0;
Matches 6; Conservative 4; Mismatches 1;
QY 2 IALRYRLAIR 12
: |||:|:|:
DB 9 LLLRYFLRLX 19
RESULT 2
LPA2_BACSU
ID LPA2_BACSU STANDARD; PRT; 24 AA.
AC P08497; P94553;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Aspartokinase II leader peptide (Attenuator peptide).
GN ASK OR AECA
OS Bacillus subtilis.
OC Bacillaria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]

34 22 37.9 36 1 PYI_RANRI
35 22 37.9 36 1 SPYI_PHYBI
36 22 37.9 37 1 PSBY_GUITH
37 22 37.9 37 1 PYI_CHICK
38 22 37.9 45 1 Y5K6_SSVI
39 22 37.9 49 1 HSP1_SAGIM
40 22 37.9 51 1 HSP1_ALOSE
41 22 37.9 54 1 RK32_TORAC
42 22 37.9 56 1 RL32_BACHD
43 22 37.9 57 1 IVBK_DENPO
44 22 37.9 58 1 FER6_METJA
45 22 37.9 59 1 YA90_ARCFU
P29204 rana ridibu
P80952 phyllomedusa
O78433 guillardia
P29203 gallus gall
P20198 sulfolobus
P24714 saguinus im
P35302 alouatta se
P12198 nicotiana t
Q9K9R2 bacillus ha
P00981 dendroaspis
O58041 methanococc
O29175 archaeoglob

RP SEQUENCE FROM N.A.
 RX MEDLINE=87250503; PubMed=3036830;
 RA Chen N.-Y., Hu F.M., Paulus H.;
 RT "Nucleotide sequence of the overlapping genes for the subunits of
 RT Bacillus subtilis aspartokinase II and their control regions";
 RL J. Biol. Chem. 262:8787-8798(1987).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=97124191; PubMed=8969504;
 RA Wipac A., Carter N., Brignell C.S., Guy J.B., Piper K.,
 RA Sanders J., Emerson P.T., Harwood C.R.;
 RT "The dnaB-pneA (256 degrees-240 degrees) region of the Bacillus
 RT subtilis chromosome containing genes responsible for stress
 RT responses, the utilization of plant cell walls and primary
 RT metabolism";
 RL Microbiology 142:3067-3078(1996).
 [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
 RA Borriss L., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
 RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.J., Comerton I.F., Cummings N.D., Daniel R.A.,
 RA Dantolof F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 RA Ertlan K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleon N.,
 RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
 RA Giuseppe G., Guy B.J., Haga K., Haich J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
 RA Joris B., Karamata D., Kasaana Y., Klaert-Blanchard M., Klein C.,
 RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Portetelle D., Potworlik S., Prescott A.M.,
 RA Pressac E., Puig P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadleir Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
 RA Sekiguchi J., Sekowska A., Seror S.J., Serro P., Shin B.S., Soldo B.,
 RA Sroczynski A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Tanakoshi A., Tanaka T., Terpstra P., Tognoni A.,
 RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
 RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
 RA Winters P., Wipac A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus
 RT subtilis";
 RL Nature 390:249-256(1997).
 [4]
 RP CHARACTERIZATION.
 RN MEDLINE=92324588; PubMed=1624109;
 RX La Y., Stevchenko T.N., Paulus H.;
 RA "Fine-structure mapping of cis-acting control sites in the lycs
 RT operon of Bacillus subtilis";
 RL FEMS Microbiol. Lett. 71:23-27(1992).
 CC -1- FUNCTION: INVOLVED IN THE REPRESSION OF THE LYSC OPERON.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: J03294; AAA87317.1; -;
 DR EMBL: J275208; CAA99579.1; -;
 DR EMBL: J299118; CAB14808.1; -;
 DR Subtilisin; BG10915; ask.

KW Leader peptide; Complete proteome.
 FT MTTAGEN 9 9 S-W: ELEVATED LEVELS OF AK II.
 FT MTTAGEN 22 22 A-G: ELEVATED LEVELS OF AK II.
 FT MTTAGEN 24 24 H-L: ELEVATED LEVELS OF AK II.
 SQ SEQUENCE 24 AA; 2832 MW; 5EA641D7B3FBC916 CRC64;
 Query Match 48.3%; Score 28; DB 1; Length 24;
 Best Local Similarity 71.4%; Pred. No. 16;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 5 RYRYLAI 11
 Db 17 RYRYLAI 23
 RESULT 3
 ID PIMS_SQUAC STANDARD; PRT; 30 AA.
 AC P82542; P83008;
 DT 15-OCT-2001 (Rel. 40, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Phospholeman-like protein (PIMS) (Fragment).
 OS Squalus acanthias (Spiny dogfish).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
 OC Elasmobranchii; Squalae; Squaloidei; Squalidae; Squalus.
 OX NCBI_TaxID=7797;
 [1]
 RP SEQUENCE.
 RC TISSUE=Rectal gland;
 RX PubMed=10961995;
 RA Mahmoud Y.A., Yorum H., Cornelius F.;
 RT "Identification of a phospholeman-like protein from shark rectal
 RT glands: evidence for indirect regulation of Na,K-ATPase by protein
 RT kinase C via a novel member of the Fxydy family";
 RL J. Biol. Chem. 275:35969-35977(2000).
 RN [2]
 RP REVISIONS TO 19; 24 AND 26.
 RC TISSUE=Rectal gland;
 RA Cornelius F., Mahmoud Y.A., Cramb G.;
 RL Submitted (DEC-2001) to the SWISS-PROT data bank.
 RN [3]
 RP SEQUENCE OF 1-25.
 RC TISSUE=Rectal gland;
 RX MEDLINE=21534268; PubMed=11676495;
 RA Schuurmans Stekhoven F.M.A.H., Flik G., Wendelaar Bonga S.B.;
 RT "terminal sequences of small ion channels in rectal glands of
 RT sharks: a biochemical hallmark for classification and phylogeny?";
 RL Biochem. Biophys. Res. Commun. 288:670-675(2001).
 CC -1- FUNCTION: Induces a hyperpolarization-activated chloride current
 CC when expressed in xenopus oocytes. May have a functional role in
 CC muscle contraction.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Mitochondrial
 CC membrane.
 CC -1- PTM: Phosphorylated by protein kinase A (PK-A) and protein kinase
 CC C (PK-C). Phosphorylated in response to insulin and adrenergic
 CC stimulation (by similarity).
 CC -1- SIMILARITY: BELONGS TO THE FXYD FAMILY.
 DR InterPro: IPR000272; ATP1G1_PLM MATR.
 DR PROSITE: PS01310; FXYD; FALSE NEG.
 KW Transmembrane, Ionic channel, Ion transport, Phosphorylation;
 KM Microsome.
 PT NON TER 30 30
 SQ SEQUENCE 30 AA; 3430 MW; 647BE1C06B2C9F1 CRC64;
 Query Match 46.6%; Score 27; DB 1; Length 30;
 Best Local Similarity 45.5%; Pred. No. 33;
 Matches 5; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
 QY 1 RIALRYRLAI 11
 Db 11 RPYDYRYRLAI 21

```

RESULT 4
DEF2_MESAU
ID DEF2_MESAU STANDARD; PRT; 31 AA.
AC P81466;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Neutrophil defensin 2 (HAMP-2).
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OC NCBI_TaxID=10036;
RN [1]
RP SEQUENCE.
RX MEDLINE=97045125; PubMed=8890190;
RA Mak P., Wojcik K., Thogersen I.B., Dubin A.;
RT "Isolation, antimicrobial activities, and primary structures of
RT hamster neutrophil defensins.";
RL Infect. Immun. 64:4444-4449(1996).
CC -!- FUNCTION: BACTERICIDAL ACTIVITY, GREATER AGAINST GRAM-POSITIVE
CC BACTERIA. LOW ANTI-FUNGI ACTIVITY.
CC -!- SIMILARITY: BELONGS TO THE CORTICOSTATIN/DEFENSIN FAMILY.
CC -!- CAUTION: THE HAMP-2 COULD BE A PRODUCT OF PROTEOLYTIC N-TERMINAL
CC AMINO ACID REMOVAL FROM HAMP-4.
DR InterPro; IPR001271; Defensin_mammal.
DR Pfam; PF00323; defensins; 1.
DR SMART; SM00048; DEFNS; 1.
DR PROSITE; PS00269; DEFENSIN; 1.
KW Defensin; Antibiotic; Fungicide.
FT DISULFID 1 29 BY SIMILARITY.
FT DISULFID 3 18 BY SIMILARITY.
FT DISULFID 8 28 BY SIMILARITY.
SQ SEQUENCE 31 AA; 3621 MW; 3DE9E747D474AD34 CRC64;

Query Match 44.8%; Score 26; DB 1; Length 31;
Best Local Similarity 41.7%; Pred. No. 54;
Matches 5; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 RIALRYVRLAIR 12
: ||| |
DB 19 RLGNFYRLCCR 30

RESULT 5
DEF4_MESAU
ID DEF4_MESAU STANDARD; PRT; 33 AA.
AC P81468;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Neutrophil defensin 4 (HAMP-4).
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OC NCBI_TaxID=10036;
RN [1]
RP SEQUENCE.
RX MEDLINE=97045125; PubMed=8890190;
RA Mak P., Wojcik K., Thogersen I.B., Dubin A.;
RT "Isolation, antimicrobial activities, and primary structures of
RT hamster neutrophil defensins.";
RL Infect. Immun. 64:4444-4449(1996).
CC -!- FUNCTION: BACTERICIDAL ACTIVITY, GREATER AGAINST GRAM-POSITIVE
CC BACTERIA. LOW ANTI-FUNGI ACTIVITY.
CC -!- SIMILARITY: BELONGS TO THE CORTICOSTATIN/DEFENSIN FAMILY.
CC -!- CAUTION: THE HAMP-2 COULD BE A PRODUCT OF PROTEOLYTIC N-TERMINAL
CC AMINO ACID REMOVAL FROM HAMP-4.
DR InterPro; IPR001271; Defensin_mammal.
DR Pfam; PF00323; defensins; 1.
DR SMART; SM00048; DEFNS; 1.

```

```

DR PROSITE; PS00269; DEFENSIN; 1.
KW Defensin; Antibiotic; Fungicide.
FT DISULFID 3 31 BY SIMILARITY.
FT DISULFID 5 20 BY SIMILARITY.
FT DISULFID 10 30 BY SIMILARITY.
SQ SEQUENCE 33 AA; 3821 MW; 23097FE7D474AD65 CRC64;

Query Match 44.8%; Score 26; DB 1; Length 33;
Best Local Similarity 41.7%; Pred. No. 58;
Matches 5; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 RIALRYVRLAIR 12
: ||| |
DB 21 RLGNFYRLCCR 32

RESULT 6
Y281_TREPA
ID Y281_TREPA STANDARD; PRT; 42 AA.
AC O83305;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein TP0281.
GN TP0281.
OS Treponema pallidum.
OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
OC NCBI_TaxID=160;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nichols;
RX MEDLINE=98332770; PubMed=9665876;
RA Fraser C.M., Norris S.J., Weinstein G.M., White O., Sutton G.G.,
RA Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,
RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
RA Khaliq H., Richardson D., Howell J.K., Chidambaram M., Uterback T.,
RA McDonald L., Artach P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA Venter J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis
RT spirochete";
RL Science 281:375-388(1998).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF001209; AAC65278.1; -
DR TIGR; TP0281; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 42 AA; 4462 MW; F9FBAD73F78BCD7F CRC64;

Query Match 44.8%; Score 26; DB 1; Length 42;
Best Local Similarity 45.5%; Pred. No. 76;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 IALRYVRLAIR 12
: ||| |
DB 17 VGLRCSRLSVR 27

RESULT 7
LCRT_YERPS
ID LCRT_YERPS STANDARD; PRT; 56 AA.
AC Q00932;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-JUL-1993 (Rel. 26, Last annotation update)
DE Low calcium response locus protein T.

```

```

GN LCRT.
OS Yersinia pseudotuberculosis.
OG Plasmid pIBL.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Yersinia.
OX NCBI_TaxID=633;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YPT11;
RX MEDLINE=92250432; PubMed=1577700;
RA Rimpplainen M., Forsberg A., Wolf-Watz H.;
RT "A novel protein, LcrQ, involved in the low-calcium response of
RL Yersinia pseudotuberculosis shows extensive homology to YopH.";
J. Bacteriol. 174:3355-3363(1992).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M83986; AAA27656.1; -.
DR InterPro: IPR001230; Prenyl_site.
KM Plasmid.
SQ SEQUENCE 56 AA; 6760 MW; E6C9AD6C40A16593 CRC64;

Query Match 43.1%; Score 25; DB 1; Length 56;
Best Local Similarity 80.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 4 LRYR 8
Db 35 LRYK 39

RESULT 8
RS21 MYCPN
ID RS21 MYCPN STANDARD; PRT; 60 AA.
AC P57079;
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 30S ribosomal protein S21.
GN RPSU OR MPN296 OR MP539.1.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelreich R., Hilbert H., Plagens H., Pirkl E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae.";
RL Nucleic Acids Res. 24:4420-4449(1996).
RN [2]
RP IDENTIFICATION.
RX MEDLINE=20411492; PubMed=10954595;
RA Dandekar T., Huynen M., Regula J.T., Ueberle B., Zimmermann C.U.,
RA Andrade M.A., Doerks T., Sanchez-Pulido L., Snel B., Suyama M.,
RA Yuan Y.P., Herrmann R., Bork P.;
RT "Re-annotating the Mycoplasma pneumoniae genome sequence: adding
RT value, function and reading frames.";
RL Nucleic Acids Res. 28:3278-3288(2000).
CC -----
CC -1- SIMILARITY: BELONGS TO THE S21P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE000052; -; NOT ANNOTATED_CDS.
DR EMBL: AE000053; AG34754.1; -.
DR InterPro: IPR001911; Ribosomal_S21.
DR PROSITE: PS01181; RIBOSOMAL_S21; FALSE_NEG.
KM Ribosomal protein, Complete proteome.
SQ SEQUENCE 60 AA; 7515 MW; 30D87DCF4635F003 CRC64;

Query Match 43.1%; Score 25; DB 1; Length 60;
Best Local Similarity 50.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 RIALRYRLAIR 12
Db 20 RVSLEIRRLAQR 31

RESULT 9
ID PSBY PORPU STANDARD; PRT; 36 AA.
AC P51206;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Photosystem II protein Y.
GN PSBY OR YCF32.
OS Porphyra purpurea.
OC Porphyra.
OC Chloroplast.
OC Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae; Porphyra.
OX NCBI_TaxID=2787;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Avonport;
RA Reith M.E., Munholland J.;
RT "Complete nucleotide sequence of the Porphyra purpurea chloroplast
RT genome.";
RL Plant Mol. Biol. Rep. 13:333-335(1995).
CC -1- FUNCTION: MANGANESE-BINDING POLYPEPTIDE WITH L-ARGININE
CC METABOLIZING ENZYME ACTIVITY. COMPONENT OF THE CORE OF PHOTOSYSTEM
CC II (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Chloroplast thylakoid membrane.
CC -1- SIMILARITY: BELONGS TO THE PSBY FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U38804; AAC08092.1; -.
KM Photosystem II; Chloroplast; Transmembrane; Thylakoid.
FT DOMAIN 1
FT TRANSMEM 6
FT DOMAIN 23
FT TRANSMEM 24
FT DOMAIN 36
FT TRANSMEM 36
SQ SEQUENCE 36 AA; 4033 MW; AE0642B60945BD23 CRC64;

Query Match 41.4%; Score 24; DB 1; Length 36;
Best Local Similarity 50.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 RIALRYR 8
Db 25 RVALQQR 32

RESULT 10
HSP1 HUMAN
ID HSP1_HUMAN STANDARD; PRT; 50 AA.

```


AC P04553;
 DT 13-AUG-1987 (Rel. 05, Created)
 DT 13-AUG-1987 (Rel. 05, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Sperm protamine P1 (Cysteine-rich protamine).
 DE PM1.
 GN Homo sapiens (Human).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=85280768; PubMed=4027356;
 RA McKay D.J., Renaux B.S., Dixon G.H.;
 RT "The amino acid sequence of human sperm protamine P1.";
 RL Biosci. Rep. 5:383-391(1985).
 RN [2]
 RP SEQUENCE.
 RX MEDLINE=86296190; PubMed=3527226;
 RA Ammer H., Henschen A., Lee C.-H.;
 RT "Isolation and amino-acid sequence analysis of human sperm protamines
 P1 and P2. Occurrence of two forms of protamine P2.";
 RL Biol. Chem. Hoppe-Seyler 367:515-522(1986).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91184796; PubMed=2081589;
 RA Domenjoud L., Nussbaum G., Adham I.M., Greeske G., Engel W.;
 RT "Genomic sequences of human protamines whose genes, PRM1 and PRM2,
 are clustered.";
 RL Genomics 8:127-133(1990).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90129065; PubMed=2613245;
 RA Krawetz S.A., Herfort M.H., Hamerton J.L., Pon R.T., Dixon G.H.;
 RT "Chromosomal localization and structure of the human P1 protamine
 gene.";
 RL Genomics 5:639-645(1989).
 RN [5]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88015622; PubMed=3658707;
 RA Lee C.-H., Hoyer-Fender S., Engel W.;
 RT "The nucleotide sequence of a human protamine 1 cDNA.";
 RL Nucleic Acids Res. 15:7639-7639(1987).
 RN [6]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95074145; PubMed=7983046;
 RA Nelson J.E., Krawetz S.A.;
 RT "Characterization of a human locus in transition.";
 RL J. Biol. Chem. 269:31067-31073(1994).
 RN [7]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94087159; PubMed=8263493;
 RA Queralt R., de Fabregues-Boixar O., Adroer R., Gene M.,
 Gomez-Catalan J., Huguet E., Oliva R.;
 RT "Direct sequencing of the human protamine P1 gene and application in
 forensic medicine.";
 RL J. Forensic Sci. 38:1491-1501(1993).
 RN [8]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20123244; PubMed=10659848;
 RA Wyckoff G.J., Wang W., Wu C.-I.;
 RT "Rapid evolution of male reproductive genes in the descent of man.";
 RL Nature 403:304-309(2000).
 RN [9]
 RP SEQUENCE FROM N.A.
 RX Tissue-Testis;
 RA Strauberg R.;
 RT Submitted (Feb-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: PROTAMINES SUBSTITUTE FOR HISTONES IN THE CHROMATIN OF
 CC SPERM DURING THE HAPLOID PHASE OF SPERMATOGENESIS. THEY COMPACT
 CC SPERM DNA INTO A HIGHLY CONDENSED, STABLE AND INACTIVE COMPLEX.
 CC -1- SUBUNIT: CROSS-LINKED BY INTERCHAIN DISULFIDE BONDS AROUND THE
 CC DNA-HELIX.

CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- TISSUE SPECIFICITY: TESTIS.
 CC -1- SIMILARITY: BELONGS TO THE PROTAMINE P1 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z46940; CAA87065.1; --
 DR EMBL; M60331; AAA63249.1; --
 DR EMBL; M29706; AAA60191.1; --
 DR EMBL; Y00443; CAA68499.1; --
 DR EMBL; U15422; AAC50486.1; --
 DR EMBL; S68144; AAB29321.2; --
 DR EMBL; AP215707; AAF34620.1; --
 DR EMBL; BC003673; AAH03673.1; --
 DR PIR; A02656; HSHUP1.
 DR PIR; A26843; A26843.
 DR PIR; A38515; A38515.
 DR PIR; A33330; A33330.
 DR PIR; A60896; A60896.
 DR Genew; HGNC:9447; PRM1.
 DR MIM; 182880; --
 DR InterPro; IPR000221; Protamine P1.
 DR PIR; A00260; protamine P1; 1.
 DR PROSITE; PS00048; PROTAMINE_P1; 1.
 KW Chromosomal protein; Nucleosome core; Spermatogenesis; DNA-binding;
 KW Testis; DNA condensation; Nuclear protein.
 FT INIT MET 0 0
 FT DISULFID 5 5 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 38 38 INTERCHAIN (WITH C-38) (BY SIMILARITY).
 FT DISULFID 39 47 BY SIMILARITY.
 SQ SEQUENCE 50 AA; 6692 MW; 4A5C13CF71B1F33D CRC64;
 Query Match 41.4%; Score 24; DB 1; Length 50;
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 5 RYR 8
 DB 13 RYR 16
 RESULT 11
 ID HSP1_PONPY STANDARD; PRT; 50 AA.
 AC P35310;
 DT 01-PEB-1994 (Rel. 28, Created)
 DT 01-PEB-1994 (Rel. 28, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Sperm protamine P1 (Cysteine-rich protamine).
 GN PRM1.
 OS Pongo pygmaeus (Orangutan).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
 OX NCBI_TaxID=9600;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94141943; PubMed=8308910;
 RA Retief J.D., Winkfein R.J., Dixon G.H., Adroer R., Queralt R.,
 RA Ballabruga J., Oliva R.;
 RT "Evolution of protamine P1 genes in primates.";
 RL J. Mol. Evol. 37:426-434(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20123244; PubMed=10659848;
 RA Wyckoff G.J., Wang W., Wu C.-I.;
 RT "Rapid evolution of male reproductive genes in the descent of man.";
 RL Nature 403:304-309(2000).

```

CC - FUNCTION: PROTAMINES SUBSTITUTE FOR HISTONES IN THE CHROMATIN OF
CC Sperm DURING THE HAPLOID PHASE OF SPERMATOGENESIS. THEY COMACT
CC Sperm DNA INTO A HIGHLY CONDENSED, STABLE AND INACTIVE COMPLEX.
CC - SUBUNIT: CROSS-LINKED BY INTERCHAIN DISULFIDE BONDS AROUND THE
CC DNA-HELIX (BY SIMILARITY).
CC - SUBCELLULAR LOCATION: Nuclear.
CC - TISSUE SPECIFICITY: TESTIS.
CC - SIMILARITY: BELONGS TO THE PROTAMINE P1 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L14589; AAA36946.1; -.
CC InterPro; IPR000221; Protamine_P1.
CC Pfam; PF00260; Protamine_P1; 1.
CC PROSITE; PS00048; PROTAMINE_P1; 1.
CC Chromosomal protein; Nucleosome core; Spermatogenesis; DNA-binding;
CC Testis; DNA condensation; Nuclear protein.
CC INT MET 0 0 BY SIMILARITY.
CC SEQUENCE 50 AA; 6660 MW; D9935BF411E5BE78 CRC64;
CC -----
Query Match 41.4%; Score 24; DB 1; Length 50;
Best Local Similarity 50.0%; Pred. No. 2.3e+02;
Matches 6; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
OY 1 RIALRYRLAIR 12
DB 37 RCCRRRYRLCR 48
-----
RESULT 12
Y331 BORBU STANDARD; PRT; 51 AA.
ID Y331 BORBU
AC O51309;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein BB0331.
GN BB0331.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
CC NCB1_TaxID=139;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35210 / B31;
R MEDLINE=98065943; PubMed=9403685;
RA Fraser C.M., Castens S., Huang W.M., Dodson R., Hickey E.K., Gwinn M.,
RA Lathigra R., White O., Ketchum K.A., Richardson D., Hanson M.,
RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA Peterson J., Keilavage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weiman J.,
RA Uterback T., Matthey L., McDonald L., Artlich P., Bowman C.,
RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochete, Borrelia
RT burgdorferi."
RT Nature 390:580-586 (1997).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE001140; AAC66728.1; -.

```

```

DR TIGR; BB0331; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 51 AA; 6444 MW; .ACBCTB55B5A8E0DA CRC64;
-----
Query Match 41.4%; Score 24; DB 1; Length 51;
Best Local Similarity 44.4%; Pred. No. 2.4e+02;
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
OY 4 LRYRLAIR 12
DB 18 LKFSYSLDR 26
-----
RESULT 13
Y072 NPVOV STANDARD; PRT; 55 AA.
ID Y072 NPVOV
AC O10325;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Hypothetical 6.3 kDa protein (ORF75).
OS Oryzia pseudotsugata multicapsid polyhedrosis virus (OpNPV).
CC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
CC Nucleopolyhedrovirus.
CC NCB1_TaxID=164623;
[1]
RP SEQUENCE FROM N.A.
RC MEDLINE=97271300; PubMed=9126251;
RA Ahrens C.H., Russell R.R., Funk C.U., Evans J., Harwood S.,
RA Rohrmann G.F.;
RT "The sequence of the Oryzia pseudotsugata multincleocapsid nuclear
RT polyhedrosis virus genome."
RT Virology 229:381-399 (1997).
CC - SIMILARITY: TO CORRESPONDING ORF IN ACNPV.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U75930; AAC59074.1; -.
CC Hypothetical protein.
KW Hypothetical protein.
SQ SEQUENCE 55 AA; 6250 MW; 64A81F5A0229AA08 CRC64;
-----
Query Match 41.4%; Score 24; DB 1; Length 55;
Best Local Similarity 50.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
OY 4 LRYRLAIR 11
DB 6 LAYVTKL 13
-----
RESULT 14
MGN CHICK STANDARD; PRT; 26 AA.
ID MGN CHICK
AC P50594;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Mago nashi protein homolog (Fragment).
GN MAGOH.
OS Gallus gallus (Chicken).
CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
CC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
CC Gallus.
CC NCB1_TaxID=9031;
[1]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart;

```

RA Shi W., Antin P., Flink I.L., Morkin E.;
RL Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- SIMILARITY: BELONGS TO THE MAGO NASHI FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L29089; -; NOT_ANNOTATED_CDS.
DR InterPro; IPR004023; Mago_nashi.
DR Pfam; PF02792; Mago_nashi; 1.
KW Nuclear protein.
FT NON TER 26 26
SQ SEQUENCE 26 AA; 3062 MW; AF82470DE9E29230 CRC64;

Query Match 39.7%; Score 23; DB 1; Length 26;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 LRYX 7
DB 7 LRYX 10

RESULT 15
PSAM_CYAPA STANDARD; PRT; 31 AA.
AC P48185;
DC 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Photosystem I reaction centre subunit XII (PSI-M).
GN PSAM.
OS Cyanophora paradoxa.
OC Cyanelle.
OC Eukaryota; Glaucocystophyceae; Cyanophoraceae; Cyanophora.
OX NCBI_TaxID=2762;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LB555 / Pringsheim;
RA Stirewalt V.L., Michalowski C.B., Loeffelhardt W., Bohnert H.J.,
RA Bryant D.A.;
RT "Nucleotide sequence of the cyanelle DNA from Cyanophora paradoxa."
RL Plant Mol. Biol. Rep. 13:327-332(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=LB555 / Pringsheim;
RA Loeffelhardt W., Stirewalt V.L., Michalowski C.B., Annarella M.,
RA Farley J.Y., Schluchter W.M., Chung S., Newmann-Spallart C.,
RA Steiner J.M., Jakowitsch J., Bohnert H.J., Bryant D.A.;
RT "The complete sequence of the cyanelle genome of Cyanophora paradoxa:
RL the genetic complexity of a primitive plastid."
RL (In) Schenk H.E.A., Herrmann R., Jeon K.W., Mueller N.E.,
RL Schwemmler W. (eds.);
RL Eukaryotism and Symbiosis, pp.40-48, Springer-Verlag, Heidelberg
RL (1997).
CC -!- SIMILARITY: BELONGS TO THE PSAM FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U30821; AAA81277.1; -;
KW Photosystem I; Photosynthesis; Cyanelle.

SQ SEQUENCE 31 AA; 3332 MW; D92CA0CE635804D9 CRC64;

Query Match 39.7%; Score 23; DB 1; Length 31;
Best Local Similarity 50.0%; Pred. No. 2.1e+02;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 RIALRYR 8
DB 23 RLALELYK 30

Search completed: December 19, 2002, 16:19:45
Job time : 6.47059 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 19, 2002, 16:16:02 ; Search time 10.4118 Seconds
(without alignments)
110.799 Million cell updates/sec

Title: US-08-653-294C-40

Perfect score: 58

Sequence: 1 RIALRYRLAIR 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 15911

Minimum DB seq length: 0

Maximum DB seq length: 60

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Pir1:*

2: Pir2:*

3: Pir3:*

4: Pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	63.8	44	4	IMBP8
2	31	53.4	33	2	AF0546
3	31	53.4	56	2	A69538
4	30	51.7	52	2	H97793
5	29	50.0	27	2	B34224
6	29	50.0	60	2	F97702
7	28	48.3	24	2	G69590
8	28	48.3	25	2	PQ0487
9	28	48.3	40	2	G82620
10	27	46.6	23	2	JS0244
11	27	46.6	27	4	I52725
12	27	46.6	45	2	D64786
13	26	44.8	42	2	D71343
14	26	44.8	43	2	S66678
15	26	44.8	52	2	D84640
16	26	44.8	53	2	A95335
17	25.5	44.0	38	2	S78728
18	25	43.1	15	2	FA0057
19	25	43.1	19	2	PH1360
20	25	43.1	23	2	PH1303
21	25	43.1	28	2	PH1335
22	25	43.1	33	2	I61695
23	25	43.1	33	2	I46596
24	25	43.1	51	2	FN0479
25	25	43.1	57	2	I370255
26	25	43.1	59	2	I38222
27	24	41.4	20	2	PH1338
28	24	41.4	31	2	A61014
29	24	41.4	32	2	A61143

30 24 41.4 32 2 D61014
31 24 41.4 32 2 E61014
32 24 41.4 36 2 S73127
33 24 41.4 36 2 A82598
34 24 41.4 36 2 C85649
35 24 41.4 37 2 B95215
36 24 41.4 39 2 G90716
37 24 41.4 39 2 G64801
38 24 41.4 48 2 D82687
39 24 41.4 50 2 F82409
40 24 41.4 51 1 HSHUP1
41 24 41.4 51 1 B70141
42 24 41.4 51 2 S51909
43 24 41.4 51 2 A84082
44 24 41.4 52 2 T31016
45 24 41.4 53 2 S01179

ALIGNMENTS

RESULT 1

IMBP8

hypothetical immunity region protein 8 - Bacillus phage phi-105

C/Species: Bacillus phage phi-105

A/Note: host Bacillus subtilis

C/Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999

C/Accession: G24521

R/Cully, D.F.; Garro, A.J.

Gene 38, 153-164, 1985

A/Title: Nucleotide sequence of the immunity region of Bacillus subtilis bacteriophage

A/Reference number: A91535; MUID:86056972; PMID:3934047

A/Accession: G24521

A/Molecule type: DNA

A/Residues: 1-44 <CUL>

A/Cross-references: GB:M11920; NID:9215477

C/Comment: This is the hypothetical translation of a sequence that was not reported as

Query Match 63.8%; Score 37; DB 4; Length 44;

Best Local Similarity 54.5%; Pred. No. 3.1;

Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 IALRYRLAIR 12

:|:|:|:|:|:

Db 9 LLLRYRLALK 19

RESULT 2

AF0546

hypothetical protein STY0391 [imported] - Salmonella enterica subsp. enterica serovar Typhimurium

C/Species: Salmonella enterica subsp. enterica serovar Typhimurium

A/Note: this species has also been called Salmonella typhimurium

C/Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 09-Nov-2001

C/Accession: AF0546

R/Farkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher

, S.; Moule, S.; O'Garra, P.

Nature 413, 848-852, 2001

A/Title: Complete genome sequence of a multiple drug resistant Salmonella enterica sero

A/Reference number: AB0502; PMID:11677608

A/Accession: AF0546

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-33 <PAR>

A/Cross-references: GB:AL513382; PIDN:CAD08814.1; PID:G16501628; GSPDB:GN00176

C/Genetics:

A/Gene: STY0391

Query Match

Best Local Similarity 53.4%; Score 31; DB 2; Length 33;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 5 RYRLAI 11
|||
Db 4 RYRPAV 10

RESULT 3

A69538 conserved hypothetical protein AF2305 - Archaeoglobus fulgidus
C/Species: Archaeoglobus fulgidus
C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
C/Accession: A69538
R/Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
R.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirsch, E.F.
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A/Authors: Overbeek, T.; Cotton, M.D.; Spriggs, T.; Artlich, P.; Kaine, B.P.; Sykes, S.
Smith, H.O.; Woese, C.R.; Venter, J.C.
A/Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaea
A/Reference number: A69538; MUID:98049343; PMID:9389475
A/Accession: A69538
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-56 <KIR>
A/Cross-references: GB:AE000945; GB:AE000782; NID:g2689268; PIDN:AA88955.1; PID:g264821

Query Match 53.4%; Score 31; DB 2; Length 56;
Best Local Similarity 70.0%; Pred. No. 53;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 3 ALRYRLAIR 12
|||
Db 32 AERYRLAIR 41

RESULT 4

H97793 hypothetical protein RC0752 [imported] - Rickettsia conorii (strain Malish 7)
C/Species: Rickettsia conorii
C/Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 30-Sep-2001
C/Accession: H97793
R/Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; R
Science 293, 2093-2098, 2001
A/Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.
A/Reference number: A97700; MUID:21442074; PMID:11557893
A/Accession: H97793
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-52 <KIR>
A/Cross-references: GB:AE006914; PIDN:AL003290.1; PID:g15619647; GSPDB:GN00173
C/Genetics:
A/Gene: RC0752

Query Match 51.7%; Score 30; DB 2; Length 52;
Best Local Similarity 50.0%; Pred. No. 76;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Oy 2 IALRYRLAI 11
|||
Db 11 LAIKYNSAI 20

RESULT 5

B34224 hypothetical protein H1.10 - chicken
N/Alternate names: histone H1.10 [misidentification]
C/Species: Gallus gallus (chicken)
C/Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 04-Mar-2000
C/Accession: B34224
R/Kirsh, A.L.; Groudine, M.; Challoner, P.B.
Genes Dev. 3, 2172-2179, 1989
A/Title: Polyadenylation and U7 snRNP-mediated cleavage: alternative modes of RNA 3' pr
A/Reference number: A91623; MUID:90185195; PMID:2576416
A/Accession: B34224

A/Status: not compared with conceptual translation
A/Molecule type: mRNA
A/Residues: 1-27 <KIR>
C/Comment: This sequence may be misidentified or contain errors.

Query Match 50.0%; Score 29; DB 2; Length 27;
Best Local Similarity 50.0%; Pred. No. 62;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Oy 2 IALRYRLAI 11
|||
Db 6 VTRHYRLAI 15

RESULT 6

F97702 hypothetical protein RC0022 [imported] - Rickettsia conorii (strain Malish 7)
C/Species: Rickettsia conorii
C/Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 30-Sep-2001
C/Accession: F97702
R/Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; Ro
Science 293, 2093-2098, 2001
A/Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.
A/Reference number: A97700; MUID:21442074; PMID:11557893
A/Accession: F97702
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-60 <KIR>
A/Cross-references: GB:AE006914; PIDN:AL002560.1; PID:g15619056; GSPDB:GN00173
C/Genetics:
A/Gene: RC0022

Query Match 50.0%; Score 29; DB 2; Length 60;
Best Local Similarity 40.0%; Pred. No. 1,4e+02;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Oy 3 ALRYRLAIR 12
|||
Db 27 SLKRYKALK 36

RESULT 7

G69590 aspartokinase II attenuator ask - Bacillus subtilis
C/Species: Bacillus subtilis
C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 15-Oct-1999
C/Accession: G69590
R/Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
A.; Ehrlich, S.D.; Emmertson, P.T.; Entian, K.D.; Errington, D.; Fabre, C.; Ferrari, E.
Nature 390, 249-256, 1997
A/Authors: Foulger, D.; Fritze, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallizi, A.; Gall
lech, U.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holzapfel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A/Authors: lauber, U.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maneel
Y., M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon,
A/Authors: Schleich, S.; Schroeter, R.; Scroferne, F.; Sekiguchi, U.; Sekowska, A.; Seror
akeuchi, M.; Yamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A/Authors: Yoshikawa, H.F.; Zunsstein, E.; Yoshikawa, H.; Danchin, A.
A/Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A/Reference number: A69580; MUID:98044033; PMID:9384377
A/Accession: G69590
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-24 <KIR>
A/Cross-references: GB:G99118; GB:AL009126; NID:g2635200; PIDN:CA814808.1; PID:el184097;
A/Experimental source: strain 168
C/Genetics:
A/Gene: ask

Query Match 48.3%; Score 28; DB 2; Length 24;

Best Local Similarity 71.4%; Pred. No. 86;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 5 RYRLAI 11
||| |||
Db 17 RYLLAV 23

RESULT 8
PQ0487
globulin 2b - taro (fragment)
C;Species: Colocasia esculenta (taro)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 17-Mar-2000
C;Accession: PQ0487
R;de Castro, L.A.B.; Carneiro, M.; de CM Neshich, D.; de Paiva, G.R.
Plant Cell 4, 1549-1559, 1992
A;Title: Spatial and temporal gene expression patterns occur during Corm development.
A;Reference number: PQ0482; MUID:93104679; PMID:1467653
A;Accession: PQ0487
A;Molecule type: protein
A;Residues: 1-25 <CAs>
C;Superfamily: plant kunitz-type proteinase inhibitor

Query Match 48.3%; Score 28; DB 2; Length 25;
Best Local Similarity 50.0%; Pred. No. 89;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 5 RYRLAIR 12
||| |||
Db 17 RYVALSR 24

RESULT 9
G82620
hypothetical protein XP1922 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C;Accession: G82620
R;anonymously, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequences
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: G82620
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-40 <SIM>
A;Cross-references: GB:AE004012; GB:AE003849; NID:g9107020; PIDN:APF84728.1; GSPDB:GN001
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, P.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carter, H
as-Neto, E.; Docena, C.; El-Dorfy, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigh
Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, H
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasek
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF1922

Query Match 48.3%; Score 28; DB 2; Length 40;
Best Local Similarity 75.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 ALRYRLA 10
||| |||
Db 30 AARYDLA 37

RESULT 10

JS0244
hypothetical 2.85K protein - liverwort (Marchantia polymorpha) chloroplast
C;Species: chloroplast Marchantia polymorpha
C;Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 18-Jun-1993
C;Accession: JS0244
R;Fukuzawa, H.; Uchida, Y.; Yamano, Y.; Ohyama, K.; Komano, T.
Agric. Biol. Chem. 49, 2725-2731, 1985
A;Title: Molecular cloning of promoters functional in Escherichia coli from chloroplast
A;Reference number: A90020
A;Accession: JS0244
A;Molecule type: DNA
A;Residues: 1-23 <FUK>
A;Cross-references: GB:X04465
C;Genetics:
A;Genome: chloroplast
C;Keywords: chloroplast

Query Match 46.6%; Score 27; DB 2; Length 23;
Best Local Similarity 62.5%; Pred. No. 1.3e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 IALRYRL 9
:||| |||
Db 1 MAIRLYRL 8

RESULT 11

152725
hypothetical BCL2/IGHV mutant fusion protein - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 21-Feb-1997
C;Accession: I52725
R;di, W.; Qu, G.Z.; Ye, P.; Zhang, X.Y.; Halabi, S.; Ehrlich, M.
Cancer Res. 55, 2876-2882, 1995
A;Title: Frequent detection of bcl-2/JH translocations in human blood and organ samples
A;Reference number: I52725; MUID:95316863; PMID:7796416
A;Accession: I52725
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-27 <RES>
A;Cross-references: GB:S78520; NID:g1042092
C;Comment: This sequence is the chimeric product of a translocation mutation.
C;Genetics:
A;Gene: BCL2/IGHV@; BCL2/JH
A;Map position: 18q21.33
C;Keywords: fusion protein

Query Match 46.6%; Score 27; DB 4; Length 27;
Best Local Similarity 63.6%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 IALRYRLAIR 12
||| ||| |||
Db 10 ILLRYRLGPR 20

RESULT 12

D64786
phage lambda ren protein homolog b0542 - Escherichia coli (strain K-12)
C;Species: Escherichia coli
C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002
C;Accession: D64786; JN0330; S36858
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A6720; MUID:97426617; PMID:9278503
A;Accession: D64786
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA

A;Residues: 1-45 <BLAT>
 A;Cross-references: GB:AE00160; GB:U00096; NID:g1786751; PIDN:AAC73643.1; PID:g1786754;
 A;Experimental source: strain K-12, Substrain MG1655
 R;Morimoto, M.; Hongo, E.; Hama-Inaba, H.; Machida, I.
 Nucleic Acids Res. 20, 3159-3165, 1992
 A;Title: Cloning and characterization of the mwc gene of Escherichia coli K-12 which co

A;Reference number: JN0329; MUID:92319648; PMID:1320256
 A;Accession: JN0330
 A;Molecule type: DNA
 A;Residues: 'MARAGILVVDGKVRV', 2-45 <MOR>
 A;Cross-references: GB:M62732
 A;Experimental source: strain K12
 C;Genetics:
 A;Map position: 12.3 min
 C;Superfamily: phage lambda ren protein

Query Match 46.6%; Score 27; DB 2; Length 45;
 Best Local Similarity 71.4%; Pred. No. 2.4e+02;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 YRRLAIR 12
 |||
 Db 2 YRFRATR 8

RESULT 13
 D71343
 hypothetical protein TP0281 - syphilis spirochete
 C;Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
 C;Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 05-Nov-1999
 C;Accession: D71343
 R;Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin
 rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Ullerdack, T.; Mcd
 they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
 Science 281, 375-388, 1998
 A;Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
 A;Reference number: A71250; MUID:98332770; PMID:9665876
 A;Accession: D71343
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-42 <COL>
 A;Cross-references: GB:AE001209; GB:AE00520; NID:g3322547; PIDN:AAC65278.1; PID:g332256
 A;Experimental source: strain Nichols
 C;Genetics:
 A;Gene: TP0281

Query Match 44.8%; Score 26; DB 2; Length 42;
 Best Local Similarity 45.5%; Pred. No. 3.5e+02;
 Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 IALRYRLAIR 12
 : || ||:
 Db 17 VGLRCSRLSVR 27

RESULT 14
 S66678
 serpin - wheat (fragment)
 C;Species: Triticum aestivum (common wheat)
 C;Date: 05-Dec-1998 #sequence_revision 05-Dec-1998 #text_change 29-Jan-1999
 C;Accession: S66678
 R;Rosenkrantz, I.; Hejgaard, J.; Rasmussen, S.K.; Bjorn, S.E.
 FEBS Lett. 343, 75-80, 1994
 A;Title: Serpins from wheat grain.
 A;Reference number: S43652; MUID:94215711; PMID:8163022
 A;Accession: S66678
 A;Status: preliminary
 A;Molecule type: protein
 A;Residues: 1-43 <ROS>
 C;Superfamily: antithrombin III

Query Match 44.8%; Score 26; DB 2; Length 43;
 Best Local Similarity 44.4%; Pred. No. 3.6e+02;

Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
 QY 1 IALRYRL 9
 : |||:
 Db 34 KVALRQPL 42

RESULT 15
 D84640
 hypothetical protein At2g24780 [imported] - Arabidopsis thaliana
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
 C;Accession: D84640
 R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; I
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanhaken, S.E.; Umayam, L.; Tallon, L.
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
 Nature 402, 761-768, 1999
 A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A;Reference number: A84420; MUID:20083487; PMID:10617197
 A;Accession: D84640
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-52 <STO>
 A;Cross-references: GB:AE002093; NID:g4559373; PIDN:AAD23034.1; GSPDB:GN00139
 C;Genetics:
 A;Gene: At2g24780
 A;Map position: 2

Query Match 44.8%; Score 26; DB 2; Length 52;
 Best Local Similarity 50.0%; Pred. No. 4.3e+02;
 Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 IALRYRLAIR 12
 |||:
 Db 12 RIRWRLRLAIR 23

Search completed: December 19, 2002, 16:22:58
 Job time : 10.4118 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 19, 2002, 16:02:26 ; Search time 23.8235 Seconds
(without alignments)
55.932 Million cell updates/sec

Title: US-08-653-294C-39
Perfect score: 31
Sequence: 1 YXXRXRLXR 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 473727

Minimum DB seq length: 0
Maximum DB seq length: 60

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : 1_ /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	23	74.2	20	16 AAR92910	HLA-B2702 CTL modu
2	23	74.2	20	16 AAR92908	HLA-B2702 CTL modu
3	23	74.2	20	19 AAW33791	Peptide B2702.84-7
4	23	74.2	20	19 AAW33793	Peptide B2702.84-7
5	22	71.0	10	19 AAW47266	Immunomodulatory p
6	22	71.0	10	19 AAW47270	Immunomodulatory p
7	22	71.0	10	22 AAY72478	Immunosuppressive
8	22	71.0	20	16 AAR95428	HLA-B2702.84-75-84
9	22	71.0	20	16 AAR95415	HLA-B7.84-75-84 Pa
10	22	71.0	20	16 AAR92909	HLA-B2702 CTL modu

11	22	71.0	20	16 AAR92911	HLA-B2702 CTL modu
12	22	71.0	20	16 AAR92913	HLA-B7 CTL modulat
13	22	71.0	20	16 AAR92907	HLA-B2702 CTL modu
14	22	71.0	20	19 AAW33778	Immunomodulating d
15	22	71.0	20	19 AAW33779	Immunomodulating d
16	22	71.0	20	19 AAW33790	Peptide B7.84-75/7
17	22	71.0	20	19 AAW33792	Peptide B2702.84-7
18	22	71.0	20	19 AAW33797	Peptide B7.84-75/7
19	21	67.7	10	19 AAW47268	Immunomodulatory p
20	21	67.7	10	19 AAW47272	Immunomodulatory p
21	19	61.3	17	19 AAW85996	CD4-variant genera
22	19	61.3	19	19 AAW85959	CD4-variant 3.. Ho
23	19	61.3	28	14 AAR34561	Domain 1 from inte
24	19	61.3	28	17 AAR97473	Chimaeric peptide
25	19	61.3	28	20 AAY09017	Domain 1 sequence
26	19	61.3	30	21 AAB36239	Human hKv1-4 N-ter
27	19	61.3	30	22 AAG13834	Novel human diagno
28	19	61.3	30	22 AAG13835	Novel human diagno
29	19	61.3	40	22 AAB74582	Cell population st
30	19	61.3	41	22 ABB40792	Peptide #8298 enco
31	19	61.3	41	22 ABB24987	Protein #6986 enco
32	19	61.3	41	22 AAM61652	Human brain expres
33	19	61.3	41	22 AAM74444	Human bone marrow
34	19	61.3	41	22 AAM34558	Human peptide enco
35	19	61.3	41	23 AAG44333	Peptide #8595 enco
36	19	61.3	42	21 AAY94651	Helix-loop-helix m
37	19	61.3	42	21 AAY94652	Helix-loop-helix m
38	19	61.3	48	23 ABB94081	Human secreted pro
39	19	61.3	49	22 AAM87238	Human immune/haema
40	19	61.3	50	20 AAY13207	Human secreted pro
41	19	61.3	51	21 AAG01170	Human secreted pro
42	19	61.3	51	23 ABP29634	Streptococcus poly
43	19	61.3	52	22 ABB66974	Drosophila melanog
44	19	61.3	54	22 ABB28855	Peptide #1506 enco
45	19	61.3	54	22 ABB34033	Peptide #1539 enco

ALIGNMENTS

RESULT 1
AAR92910
ID AAR92910 standard; peptide; 20 AA.
XX
AC AAR92910;
XX
DT 16-MAY-1996 (first entry)
XX
DE HLA-B2702 CTL modulating peptide (B2702.84-75(T)/75-84(T)).
XX
KW Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;
KW immunosuppressant; graft versus host disorder; transplantation; therapy;
KW class I MHC; HLA-B2702.
XX
OS Synthetic.
XX
PN W09526979-A1.
XX
PD 12-OCT-1995.
XX
PF 05-APR-1995; 95WO-US04349.
XX
PR 05-APR-1994; 94US-0222851.
XX
XX (STRD) UNIV LELAND STANFORD JUNIOR.
XX
PI Clayberger C, Krensky AM, Parham P;
XX
DR WPI; 1995-358582/46.
XX
PT Extension of acceptance period of transplants from MHC unmatched
PT donor hosts - using Class I B75-84 MHC antigen of the recipient
PT host

XX Example 15; Page 36; 80pp; English.

PS
XX
XX AAR83061-R83085, AAR83090-R83096 and AAR92907-R92913 represent fragments
CC of class I major histocompatibility complex (MHC) antigens. This
CC sequence is an inverted dimer of residues 75-84 of the alpha-1 domain of
CC the class I MHC HLA-B*2702. These sequences can be used to extend the
CC period of acceptance by a recipient of a transplant from an MHC unmatched
CC donor. The peptides are administered to a patient in conjunction with a
CC subtherapeutic amount of an immunosuppressant. This is administered to
CC the patient for a limited period of time (compared to the lifetime
CC administration for current treatments). The peptides particularly
CC modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)
CC of the patient.

CC
XX
XX Sequence 20 AA;

Query Match 74.2%; Score 23; DB 16; Length 20;
Best Local Similarity 50.0%; Pred. No. 9;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Oy 1 YXXXXRLXER 10
1 YRLATRLNER 10

Db 1 YRLATRLNER 10

RESULT 2
AAR92908
ID AAR92908 standard; peptide; 20 AA.
XX
XX AAR92908;
XX
XX 16-MAY-1996 (first entry)
XX
XX HLA-B*2702 CTL modulating peptide (B2702.84-75(T)/75-84).
XX
XX Cytotoxic T lymphocyte, CTL; major histocompatibility complex; MHC;
XX immunosuppressant; graft versus host disorder; transplantation; therapy;
XX class I MHC; HLA-B*2702.
XX
XX Synthetic.
XX
XX WO9526979-A1.
XX
XX 12-OCT-1995.
XX
XX 05-APR-1995; 95WO-US04349.
XX
XX 05-APR-1994; 94US-0222851.
XX
XX (STRD) UNIV LELAND STANFORD JUNIOR.
XX
XX Clayberger C, Krensky AM, Parham P;
XX
XX WPI; 1995-358582/46.
XX
XX Extension of acceptance period of transplants from MHC unmatched
XX donor hosts - using Class I B*75-84 MHC antigen of the recipient
XX host

PS Example 15; Page 36; 80pp; English.

XX
XX AAR83061-R83085, AAR83090-R83096 and AAR92907-R92914 represent fragments
CC of class I major histocompatibility complex (MHC) antigens. This
CC sequence is an inverted dimer of residues 75-84 of the alpha-1 domain of
CC the class I MHC HLA-B*2702. These sequences can be used to extend the
CC period of acceptance by a recipient of a transplant from an MHC unmatched
CC donor. The peptides are administered to a patient in conjunction with a
CC subtherapeutic amount of an immunosuppressant. This is administered to
CC the patient for a limited period of time (compared to the lifetime
CC administration for current treatments). The peptides particularly
CC modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)
CC of the patient.

XX
XX Sequence 20 AA;

Query Match 74.2%; Score 23; DB 16; Length 20;
Best Local Similarity 50.0%; Pred. No. 9;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Oy 1 YXXXXRLXER 10
1 YRLATRLNER 10

Db 1 YRLATRLNER 10

RESULT 3
AAW33791
ID AAW33791 standard; peptide; 20 AA.
XX
XX AAW33791;
XX
XX 19-JUN-1998 (first entry)
XX
XX Peptide B2702.84-75(T)/75-84 tested for immunomodulating activity.
XX
XX Immunomodulating dimer; immunosuppressant drug; CTL activation;
XX transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;
XX rejection.
XX
XX Synthetic.
XX
XX Homo.sapiens.
XX
XX WO9744351-A1.
XX
XX 27-NOV-1997.
XX
XX 22-MAY-1997; 97WO-US08689.
XX
XX 24-MAY-1996; 96US-0653294.
XX
XX (STRD) UNIV LELAND STANFORD JUNIOR.
XX
XX Beulow R, Clayberger C, Krensky AM;
XX
XX WPI; 1998-086530/08.
XX
XX New immunomodulating dimer peptide(s) - based on a Class I HLA-B
XX alpha-1 domain, used for preventing rejection of transplants or
XX treating autoimmune diseases

PS Example 1; Page 19; 41pp; English.

XX
XX Peptides AAW33784-98 and AAW33778-9 were assayed for their
CC immunomodulating activity. A peptide-type compound or variant is claimed
CC which has immunomodulating activity, including the N-terminal acylated
CC and/or C-terminal amidated or esterified forms of up to 60 amino acids,
CC where the peptide-type compound comprises the formula: A-B, where A, B =
CC (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or V, aa77 =
CC D, S or N, aa79 = R or G, aa80 = I or N, aa81, aa84 = a hydrophobic or
CC small amino acid; aa82 = R or L; aa83 = G or R; and aa represents amino
CC acid. The sequence in the brackets may optionally be absent or truncated
CC at any peptide type bond within the brackets. The compounds comprise
CC amino acid sequences related to a Class I HLA-B alpha1 domain (positions
CC 79-84). They can be used to inhibit cytotoxic T-lymphocytes (CTL) from
CC undesirably attacking cells in a host or in vitro. They can also be
CC used in combination with antigenic peptides or proteins of interest to
CC activate CTLs. They can also inhibit the proliferation of T cells in
CC response to anti-CD3. The peptide can be used for preventing rejection
CC of transplants or for treating autoimmune diseases, e.g. diabetes,
CC rheumatoid arthritis and lupus erythematosus. The products can also be
CC used for detection and diagnosis.

XX
XX Sequence 20 AA;

Query Match 74.2%; Score 23; DB 19; Length 20;
Best Local Similarity 50.0%; Pred. No. 9;

Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 YXXXXRLXER 10
| | | | |
Db 1 YRLATRLNER 10

RESULT 4

AAW33793
ID AAW33793 standard; peptide; 20 AA.

XX AC AAW33793;

XX DT 19-JUN-1998 (first entry)

XX DE Peptide B2702.84-75T/75-84T tested for immunomodulating activity.

XX KW Immunomodulating dimer; immunosuppressant drug; CTL activation;

XX KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain; rejection.

XX OS Synthetic.

XX OS Homo-sapiens.

XX PN W09744351-A1.

XX PD 27-NOV-1997.

XX PF 22-MAY-1997; 97WO-US08689.

XX PR 24-MAY-1996; 96US-0653294.

XX PA (STRD) UNIV LELAND STANFORD JUNIOR.

XX PI Beulow R, Clayberger C, Krensky AM;

XX DR WPI; 1998-086530/08.

XX PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B alpha-1 domain, used for preventing rejection of transplants or treating autoimmune diseases

XX PS Example 1; Page 19; 41pp; English.

XX CC Peptides AAW33784-98 and AAW33778-9 were assayed for their immunomodulating activity. A peptide-type compound or variant is claimed which has immunomodulating activity, including the N-terminal acylated and/or C-terminal amidated or esterified forms of up to 60 amino acids, where the peptide-type compound comprises the formula: A-B, where A, B = (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or V; aa77 = D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a hydrophobic or small amino acid; aa82 = R or L; aa83 = G or R; and aa represents amino acid. The sequence in the brackets may optionally be absent or truncated at any peptide type bond within the brackets. The compounds comprise amino acid sequences related to a Class I HLA-B alpha-1 domain (positions 79-84). They can be used to inhibit cytotoxic T-lymphocytes (CTL) from undesirably attacking cells in a host or in vitro. They can also be used in combination with antigenic peptides or proteins of interest to activate CTLs. They can also inhibit the proliferation of T cells in response to anti-CD3. The peptide can be used for preventing rejection of transplants or for treating autoimmune diseases, e.g. diabetes, rheumatoid arthritis and lupus erythematosus. The products can also be used for detection and diagnosis.

XX SQ Sequence 20 AA;

Query Match 74.2%; Score 23; DB 19; Length 20;

Best Local Similarity 50.0%; Pred. No. 9;

Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 YXXXXRLXER 10

| | | | |

Db 1 YRLATRLNER 10

RESULT 5

AAW47266

ID AAW47266 standard; peptide; 10 AA.

XX AC AAW47266;

XX DT 22-MAY-1998 (first entry)

XX DE Immunomodulatory peptide.

XX KW Immunomodulator; Class I HLA-B alpha-1 domain; inhibition; transplant rejection; treatment; autoimmune disease.

XX OS Homo sapiens.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT Misc-difference 1..10

FT /notes "at least one of the amino acids is the D-isomer"

XX PN W09744052-A1.

XX PD 27-NOV-1997.

XX PF 23-APR-1997; 97WO-US06705.

XX PR 22-MAY-1996; 96US-0651650.

XX PA (STRD) UNIV LELAND STANFORD JUNIOR.

XX PI Clayberger C, Krensky AM;

XX DR WPI; 1998-018220/02.

XX PT Novel immunomodulatory peptide-type compound - useful for inhibiting transplant rejection

XX PS Claim 10; Page 36; 41pp; English.

XX CC The present sequence is an immunomodulatory peptide, which comprises a Class I HLA-B alpha-1 domain sequence. It can be used in a pharmaceutical composition together with a subtherapeutic dose of an immunosuppressant, to extend the period of acceptance of a transplant from a major histocompatibility complex (MHC) unmatched donor, i.e. to inhibit transplant rejection. It can also be used in the treatment of autoimmune diseases.

XX CC Peptides using the D-form amino acids are more effective immunomodulators than their diastereomers or enantiomers.

XX SQ Sequence 10 AA;

Query Match 71.0%; Score 22; DB 19; Length 10;

Best Local Similarity 50.0%; Pred. No. 8.9;

Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 YXXXXRLXER 10

| | | | |

Db 1 YRLATRLNER 10

RESULT 6

AAW47270

ID AAW47270 standard; peptide; 10 AA.

XX AC AAW47270;

XX DT 22-MAY-1998 (first entry)

XX KW Immunomodulatory peptide.

XX DE

KW Immunomodulator; Class I HLA-B alpha-1 domain; inhibition;
 KW transplant rejection; treatment; autoimmune disease.
 XX Homo sapiens.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 1..10
 FT /note= "at least one of the amino acids is the
 FT D-isomer"
 XX
 PN MO9744052-A1.
 XX
 PD 27-NOV-1997.
 XX
 PP 23-APR-1997; 97WO-US06705.
 XX
 PR 22-MAY-1996; 96US-0651650.
 XX
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 XX
 PI Clayberger C, Krensky AM;
 XX
 DR WPI; 1998-018220/02.
 XX
 PT Novel immunomodulatory peptide-type compound - useful for inhibiting
 PT transplant rejection
 PS Claim 10; Page 36; 41pp; English.
 XX
 CC The present sequence is an immunomodulatory peptide, which
 CC comprises a Class I HLA-B alpha-1 domain sequence. It can be used
 CC in a pharmaceutical composition together with a subtherapeutic dose
 CC of an immunosuppressant, to extend the period of acceptance of a
 CC transplant from a major histocompatibility complex (MHC) unmatched
 CC donor i.e. to inhibit transplant rejection. It can also be used in
 CC the treatment of autoimmune diseases.
 CC Peptides using the D-form amino acids are more effective
 CC immunomodulators than their diastereomers or enantiomers.
 CC
 SQ Sequence 10 AA;
 XX
 QY Query Match 71.0%; Score 22; DB 19; Length 10;
 Best Local Similarity 50.0%; Pred. No. 8.9;
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 XX
 DB 1 YXXXXRLXER 10
 1 YRLAIRLDER 10
 XX
 RESULT 7
 ID AAY72478 standard; peptide; 10 AA.
 XX
 AC AAY72478;
 XX
 DT 24-APR-2001 (first entry)
 XX
 DE Immunosuppressive peptide, 2702.84-75, to prevent allograft rejection.
 XX
 KW Immunosuppressive; allograft rejection; topological parameter;
 KW physico-chemical parameter; in silico screening; pharmaceutical;
 KW cosmetic; agrochemical; biomaterial; veterinary application.
 XX
 OS Unidentified.
 XX
 PN WO200079263-A2.
 XX
 PD 28-DEC-2000.
 XX
 PI 15-MAY-2000; 2000WO-EP04338.
 XX

PR 18-JUN-1999; 99EP-0401526.
 XX
 PA (SYNT-) SYNT:EM SA.
 XX
 PI Lahana R, Clair P, Yasri A;
 XX
 DR WPI; 2001-091623/10.
 XX
 PT Identifying active candidate molecules on the basis of selected
 PT physico-chemical parameters, for in silico screening of compounds
 PT useful in pharmaceuticals, cosmetics, veterinary applications and
 PT agrochemicals -
 XX
 PS Disclosure; Page 38; 55pp; English.
 XX
 CC The present invention relates to a method for identifying
 CC physico-chemical and/or topological parameters associated with biological
 CC activity. The method involves selecting the first subset from
 CC predetermined set of physico-chemical parameters, determining their value
 CC of function, and selecting the second subset from physico-chemical
 CC parameters based on the values of function, such that each second subset
 CC is more closely associated with the activity than the first subset. The
 CC selected physico-chemical parameters are useful for developing criteria
 CC for screening candidate molecules and are suitable for use in silico
 CC screening of compounds. The compounds may be used in pharmaceuticals,
 CC cosmetics, agrochemicals, biomaterials and in veterinary applications. It
 CC is also useful as an antibiotic or antifungal agent. The present
 CC sequence is an immunosuppressive peptide, 2702.84-75 (also referred
 CC as D2702.75-84), identified by in silico screening. The immunosuppressive
 CC activity of the peptide that prevents allograft rejection is tested in a
 CC heterotopic allograft model of mouse.
 XX
 SQ Sequence 10 AA;
 XX
 QY Query Match 71.0%; Score 22; DB 22; Length 10;
 Best Local Similarity 50.0%; Pred. No. 8.9;
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 XX
 DB 1 YXXXXRLXER 10
 1 YRLAIRLNER 10
 XX
 RESULT 8
 ID AAR95428 standard; peptide; 20 AA.
 XX
 AC AAR95428;
 XX
 DT 12-NOV-1996 (first entry)
 XX
 DE HLA-B2702 84-75-84 palindrome.
 XX
 KW HLA; p74; alpha1-helix; human-leucocyte-associated antigen; inhibitor;
 KW T-cell lyase; membrane protein; mammal; heat shock protein; Hsc70; APC;
 KW B cell; calcium influx; cytotoxic T lymphocyte; CTL; differentiation;
 KW cytolysis; antigen presenting cell.
 XX
 OS Synthetic.
 XX
 PN WO9513288-A1.
 XX
 PD 18-MAY-1995.
 XX
 PF 10-NOV-1994; 94WO-US12985.
 XX
 PR 10-NOV-1993; 93US-0150493.
 XX
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 XX
 PI Clayberger C, Krensky AM;
 XX
 DR WPI; 1995-194027/25.
 XX

XX Compns. comprising lymphoid surface membrane proteins - which may
PT inhibit cytolytic activity and differentiation of CTLs.
XX
PS Example; Page 12; 29pp; English.
XX
CC AAR95413, and AAR95415-R95431 represent palindromes and fragments of
CC human-leucocyte-associated antigens. This sequence represents the
CC HLA-B2702 84-75-84 palindrome. These sequences can be used to isolate
CC the protein p74 from a T-cell lysate. p74 is a T-cell surface membrane
CC protein associated with T-cell activation in mammalian T-cells, and is
CC also immunologically cross reactive with the heat shock protein Hsc70.
CC p74 is found in a limited number of cell types, but is particularly
CC expressed on B and T cells. p74 can be isolated by lysis of a suitable
CC cell with an amphoteric detergent, and then passed through an affinity
CC column containing a covalently bound HLA-B2702 palindromic peptide.
CC Compositions comprising the extracellular fragment of p74 combined with
CC HLA-B2702.60-84 (see AAR95416), induces calcium influx, and inhibits
CC cytotoxic T lymphocyte (CTL) differentiation or cytotoxicity. Candidate
CC compounds can be screened for their effect on the cytolytic activity of
CC T-cells, by combining them with the extracellular portion of p74 and
CC determining the amount of binding between the candidate compound and
CC p74. Modulation of CTL activity can be inhibited in a cellular
CC composition containing T-cells and antigen presenting cells (APCs), by
CC adding to the mix the extracellular portion of p74, in an amount
CC sufficient to compete with p74 for the binding of the p74 ligand.
XX
SQ Sequence 20 AA;
Query Match 71.0%; Score 22; DB 16; Length 20;
Best Local Similarity 50.0%; Pred. No. 18;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 YXXXXRLXER 10
| | | | |
Db 1 YRLAIRLNER 10

RESULT 9
AAR95415
ID AAR95415 standard; peptide; 20 AA.
AC
XX AAR95415;
XX
DT 12-NOV-1996 (first entry)
XX
XX HLA-B7.84-75-84 Palindrome.
XX
XX HLA; p74; alpha-helix; human-leucocyte-associated antigen; inhibitor;
XX T-cell lysate; membrane protein; mammal; heat shock protein; Hsc70; APC;
XX B cell; calcium influx; cytotoxic T lymphocyte; CTL; differentiation;
XX cytotoxicity; antigen presenting cell.
XX
XX Synthetic.
XX
XX WO9513288-A1.
XX
XX 18-MAY-1995.
XX
XX 10-NOV-1994; 94WO-US12985.
XX
XX 10-NOV-1993; 93US-0150493.
XX
XX (STRD) UNIV LELAND STANFORD JUNIOR.
XX
XX Clayberger C, Krensky AM;
XX
XX WPI; 1995-194027/25.
XX
XX Compns. comprising lymphoid surface membrane proteins - which may
PT inhibit cytolytic activity and differentiation of CTLs.
XX
XX Example; Page 18; 29pp; English.

XX AAR95413, and AAR95415-R95431 represent palindromes and fragments of
CC human-leucocyte-associated antigens. This sequence represents the
CC HLA-B7.84-75-84 palindrome. These sequences can be used to isolate
CC the protein p74 from a T-cell lysate. p74 is a T-cell surface membrane
CC protein associated with T-cell activation in mammalian T-cells, and is
CC also immunologically cross reactive with the heat shock protein Hsc70.
CC p74 is found in a limited number of cell types, but is particularly
CC expressed on B and T cells. p74 can be isolated by lysis of a suitable
CC cell with an amphoteric detergent, and then passed through an affinity
CC column containing a covalently bound HLA-B2702 palindromic peptide.
CC Compositions comprising the extracellular fragment of p74 combined with
CC HLA-B2702.60-84 (see AAR95416), induces calcium influx, and inhibits
CC cytotoxic T lymphocyte (CTL) differentiation or cytotoxicity. Candidate
CC compounds can be screened for their effect on the cytolytic activity of
CC T-cells, by combining them with the extracellular portion of p74 and
CC determining the amount of binding between the candidate compound and
CC p74. Modulation of CTL activity can be inhibited in a cellular
CC composition containing T-cells and antigen presenting cells (APCs), by
CC adding to the mix the extracellular portion of p74, in an amount
CC sufficient to compete with p74 for the binding of the p74 ligand.
XX
SQ Sequence 20 AA;
Query Match 71.0%; Score 22; DB 16; Length 20;
Best Local Similarity 50.0%; Pred. No. 18;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 YXXXXRLXER 10
| | | | |
Db 1 YGLRLRLSER 10

RESULT 10
AAR92909
ID AAR92909 standard; peptide; 20 AA.
XX
AC AAR92909;
XX
XX 16-MAY-1996 (first entry)
XX
XX HLA-B2702 CTL modulating peptide (B2702.84-75/75-84(T)).
XX
XX Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;
XX immunosuppressant; graft versus host disorder; transplantation; therapy;
XX class I MHC; HLA-B2702.
XX
XX Synthetic.
XX
XX WO9526979-A1.
XX
XX 12-OCT-1995.
XX
XX 05-APR-1995; 95WO-US04349.
XX
XX 05-APR-1994; 94US-0222851.
XX
XX (STRD) UNIV LELAND STANFORD JUNIOR.
XX
XX Clayberger C, Krensky AM, Parham P;
XX
XX WPI; 1995-358582/46.
XX
XX Extension of acceptance period of transplants from MHC unmatched
PT donor hosts - using Class I B75-84 MHC antigen of the recipient
PT host
XX
XX Example 15; Page 36; 80pp; English.
XX
XX AAR93061-R93085, AAR93090-R93096 and AAR92907-R92913 represent fragments
CC of class I major histocompatibility complex (MHC) antigens. This
CC sequence is an inverted dimer of residues 75-84 of the alpha-1 domain of
CC the class I MHC HLA-B2702. These sequences can be used to extend the

CC period of acceptance by a recipient of a transplant from an MHC unmatched
 CC donor. The peptides are administered to a patient in conjunction with a
 CC subtherapeutic amount of an immunosuppressant. This is administered to
 CC the patient for a limited period of time (compared to the lifetime
 CC administration for current treatments). The peptides particularly
 CC modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)
 CC of the patient.
 CC
 SQ Sequence 20 AA;
 Query Match 71.0%; Score 22; DB 16; Length 20;
 Best Local Similarity 50.0%; Pred. No. 18;
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 OY 1 YXXXXRLXER 10
 DB 1 YRLAIRLNER 10
 RESULT 11
 AAR92911
 ID AAR92911 standard; peptide; 20 AA.
 XX
 AC AAR92911;
 DT 16-MAY-1996 (first entry)
 XX
 DE HLA-B2702 CTL modulating peptide (B2702.84-75/84-75).
 XX
 KW Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;
 KW immunosuppressant; graft versus host disorder; transplantation; therapy;
 KW class I MHC; HLA-B2702.
 XX
 OS Synthetic.
 XX
 PN WO9526979-A1.
 PD 12-OCT-1995.
 XX
 PF 05-APR-1995; 95WO-US04349.
 XX
 PR 05-APR-1994; 94US-0222851.
 XX
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 XX
 PI Clayberger C, Krensky AM, Parham P;
 XX
 DR WPI; 1995-358582/46.
 XX
 PT Extension of acceptance period of transplants from MHC unmatched
 PT donor hosts - using Class I B75-84 MHC antigen of the recipient
 PT host
 XX
 PS Example 15; Page 36; 80pp; English.
 XX
 CC AAR83061-R83085, AAR83090-R83096 and AAR92907-R92914 represent fragments
 CC of class I major histocompatibility complex (MHC) antigens. This
 CC sequence is a dimer of residues 75-84 of the alpha-1 domain of the class
 CC I MHC HLA-B2702. These sequences can be used to extend the period of
 CC acceptance by a recipient of a transplant from an MHC unmatched donor.
 CC The peptides are administered to a patient in conjunction with a
 CC subtherapeutic amount of an immunosuppressant. This is administered to
 CC the patient for a limited period of time (compared to the lifetime
 CC administration for current treatments). The peptides particularly
 CC modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)
 CC of the patient.
 CC
 SQ Sequence 20 AA;
 Query Match 71.0%; Score 22; DB 16; Length 20;
 Best Local Similarity 50.0%; Pred. No. 18;
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 YXXXXRLXER 10
 DB 1 YRLAIRLNER 10
 RESULT 12
 AAR92913
 ID AAR92913 standard; peptide; 20 AA.
 XX
 AC AAR92913;
 DT 16-MAY-1996 (first entry)
 XX
 DE HLA-B7 CTL modulating peptide (B7.84-75/75-84).
 XX
 KW Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;
 KW immunosuppressant; graft versus host disorder; transplantation; therapy;
 KW class I MHC; HLA-B7.
 XX
 OS Synthetic.
 XX
 PN WO9526979-A1.
 PD 12-OCT-1995.
 XX
 PF 05-APR-1995; 95WO-US04349.
 XX
 PR 05-APR-1994; 94US-0222851.
 XX
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 XX
 PI Clayberger C, Krensky AM, Parham P;
 XX
 DR WPI; 1995-358582/46.
 XX
 PT Extension of acceptance period of transplants from MHC unmatched
 PT donor hosts - using Class I B75-84 MHC antigen of the recipient
 PT host
 XX
 PS Example 15; Page 36; 80pp; English.
 XX
 CC AAR83061-R83085, AAR83090-R83096 and AAR92907-R92913 represent fragments
 CC of class I major histocompatibility complex (MHC) antigens. This
 CC sequence is an inverted dimer of residues 75-84 of the alpha-1 domain of
 CC the class I MHC HLA-B7. These sequences can be used to extend the period
 CC of acceptance by a recipient of a transplant from an MHC unmatched
 CC donor. The peptides are administered to a patient in conjunction with a
 CC subtherapeutic amount of an immunosuppressant. This is administered to
 CC the patient for a limited period of time (compared to the lifetime
 CC administration for current treatments). The peptides particularly
 CC modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)
 CC of the patient.
 CC
 SQ Sequence 20 AA;
 Query Match 71.0%; Score 22; DB 16; Length 20;
 Best Local Similarity 50.0%; Pred. No. 18;
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 OY 1 YXXXXRLXER 10
 DB 1 YGRILRLSER 10
 RESULT 13
 AAR92907
 ID AAR92907 standard; peptide; 20 AA.
 XX
 AC AAR92907;
 DT 16-MAY-1996 (first entry)
 XX
 DE HLA-B2702 CTL modulating peptide (B2702.84-75/75-84).

100

XX
DR WPI: 1998-086530/08.

XX
PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B
PT alpha-1 domain, used for preventing rejection of transplants or
PT treating autoimmune diseases

XX
PS Claim 16; Page 35; 41pp; English.

CC This sequence represents a specifically claimed immunomodulating
CC dimer peptide of the invention. A peptide-type compound or variant is
CC claimed which has immunomodulating activity, including the N-terminal
CC acylated and/or C-terminal amidated or esterified forms of up to 60
CC amino acids, where the peptide-type compound comprises the formula; A-B,
CC where A, B = (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or
CC V; aa77 = D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a
CC hydrophobic or small amino acid; aa82 = R or L; aa83 = G or R; and aa
CC represents amino acid. The sequence in the brackets may optionally be
CC absent or truncated at any peptide type bond within the brackets. The
CC compounds comprise amino acid sequences related to a Class I HLA-B
CC alpha1 domain (positions 79-84). They can be used to inhibit cytotoxic
CC T-lymphocytes (CTL) from undesirably attacking cells in a host or in
CC vitro. They can also be used in combination with antigenic peptides or
CC proteins of interest to activate CTLs. They can also inhibit the
CC proliferation of T cells in response to anti-CD3. The peptide can be
CC used for preventing rejection of transplants or for treating autoimmune
CC diseases, e.g. diabetes, rheumatoid arthritis and lupus erythematosus.
CC The products can also be used for detection and diagnosis.

XX
SQ Sequence 20 AA;

Query Match 71.0%; Score 22; DB 19; Length 20;
Best Local Similarity 50.0%; Pred. No. 18;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Oy 1 YXXXXRLXER 10
Db 1 YRLAIRLNER 10

Search completed: December 19, 2002, 16:19:06
Job time : 24.8235 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 19, 2002, 16:14:32 ; Search time 17.6471 Seconds
(without alignments)
116.760 Million cell updates/sec

Title: US-08-653-294C-39
Perfect score: 31
Sequence: 1 YXXXXLXER 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 45785

Minimum DB seq length: 0
Maximum DB seq length: 60

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL 21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	23	74.2	56	16	Q99UI6
2	20	64.5	39	4	Q9UBN9
3	19	61.3	24	5	Q94374
4	19	61.3	25	11	Q63988
5	19	61.3	38	5	Q9NKY0
6	19	61.3	52	5	Q9VEZ8
7	19	61.3	53	16	Q8UH70
8	19	61.3	56	17	Q27979
9	19	61.3	58	5	Q9Y188
10	19	61.3	58	16	Q98M32
11	19	61.3	58	16	Q925Z8
12	18	58.1	25	5	Q9BM19
13	18	58.1	26	5	Q9BM11
14	18	58.1	28	5	Q9BM76
15	18	58.1	28	5	Q9BM75
16	18	58.1	28	5	Q9BM74

17	18	58.1	28	12	Q86528	Q86528 grapevine f
18	18	58.1	38	5	Q96617	Q96617 echinococcu
19	18	58.1	46	16	Q982H9	Q982H9 rhizobium 1
20	18	58.1	48	8	Q36270	Q36270 zea mays (m
21	18	58.1	48	8	Q36278	Q36278 zea mays (m
22	18	58.1	48	17	Q9HR17	Q9HR17 halobacteri
23	18	58.1	52	7	Q31325	Q31325 aphelocoma
24	18	58.1	52	7	Q31328	Q31328 aphelocoma
25	18	58.1	52	7	Q31329	Q31329 aphelocoma
26	18	58.1	52	7	Q31330	Q31330 aphelocoma
27	18	58.1	52	7	Q31332	Q31332 aphelocoma
28	18	58.1	53	7	Q8SNE5	Q8SNE5 gallinago m
29	18	58.1	53	7	Q8SNE2	Q8SNE2 gallinago m
30	18	58.1	53	12	Q86959	Q86959 human papil
31	18	58.1	53	16	Q98FU0	Q98FU0 rhizobium 1
32	18	58.1	54	13	Q91873	Q91873 geospiza sc
33	18	58.1	54	13	Q91AM9	Q91AM9 geospiza sc
34	18	58.1	54	13	Q91ANO	Q91ANO geospiza sc
35	18	58.1	54	13	Q91ANI	Q91ANI geospiza sc
36	18	58.1	54	13	Q91AN2	Q91AN2 geospiza sc
37	18	58.1	54	13	Q91AN4	Q91AN4 geospiza sc
38	18	58.1	54	13	Q91AN5	Q91AN5 geospiza sc
39	18	58.1	54	13	Q91AN6	Q91AN6 geospiza sc
40	18	58.1	54	16	Q92P27	Q92P27 rhizobium m
41	18	58.1	55	7	Q8WM92	Q8WM92 cactospiza
42	18	58.1	55	7	Q8WM91	Q8WM91 cactospiza
43	18	58.1	55	7	Q8WM90	Q8WM90 certhidea o
44	18	58.1	55	7	Q8WM89	Q8WM89 platyspiza
45	18	58.1	55	7	Q8WM88	Q8WM88 melanospiza

ALIGNMENTS

RESULT 1

Q99UI6 PRELIMINARY; PRT; 56 AA.
AC Q99UI6;
DT 01-JUN-2001 (TREMREL. 17, Created)
DT 01-JUN-2001 (TREMREL. 17, Last sequence update)
DT 01-WAR-2002 (TREMREL. 20, Last annotation update)
DE Hypothetical protein SAV1287.
GN SAV1287 OR SAI129.1 OR SAS037.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699), and
OC Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus
OX NCBI_TaxID=158678, 158879;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S.aureus (strain Mu50), and S.aureus (strain N315);
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.,
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus.";
RL Lancet 357:1225-1240(2001).
DR EMBL; AP003361; BAB57449.1; --
DR EMBL; AP003133; BAB42382.1; --
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 56 AA; 6474 MW; FADB3071F92F683D CRC64;

Query Match 74.2%; Score 23; DB 16; Length 56;
Best Local Similarity 50.0%; Pred. No. 43;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 YXXXXLXER 10
|
|
|
|

Db 21 YINAYRLAER 30

RESULT 2

Q9UBN9 PRELIMINARY; PRT; 39 AA.
 AC Q9UBN9;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE E6-AP ubiquitin-protein ligase (Fragment).
 GN UBE3A.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OK NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=BRAIN;
 RX MEDLINE=98126441; PubMed=9465301;
 RA Kishino T., Magstaf J.;
 RT "Genomic organization of the UBE3A/E6-AP gene and related
 pseudogenes.";
 RL Genomics 47:101-107(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Hennies H.C., Buerger J., Sperling K., Reis A.;
 RT "Mutations in the E6-AP gene (UBE3A) in patients with Angelman
 syndrome.";
 RL Submitted (AUG-1997) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AF009341; AAC39580.1; -;
 DR EMBL; AJ001113; CA04540.1; -;
 DR InterPro; IPR000569; HECT_domain.
 DR Pfam; PF00632; HECT, 1.
 DR PROSITE; PS50237; HECT, 1.
 KW Ligase.
 FT NON TER
 SQ SEQUENCE 39 AA; 4441 MW; 763722F374FA7193 CRC64;
 Query Match 64.5%; Score 20; DB 4; Length 39;
 Best Local Similarity 40.0%; Pred. No. 2.1e+02;
 Matches 4; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 YXXXXRLXER 10
 Db 16 YSSKRLKER 25

RESULT 3
 Q94374 PRELIMINARY; PRT; 24 AA.
 AC Q94374;
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Pp1-like Ser/Thr protein phosphatase (Fragment).
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditidae; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OK NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Zeke T., Gergely P., Dombradi V.;
 RT "The catalytic subunits of Ser/Thr protein phosphatases from
 Caenorhabditis elegans: A biochemical and molecular biological
 survey.";
 RL Submitted (JUL-1996) to the EMBL/Genbank/DBJ databases.
 CC -I- CATALYTIC ACTIVITY: A PHOSPHOROTRANSFERIN + H(2)O = A PROTEIN +
 PHOSPHATE.
 CC -I- SIMILARITY: BELONGS TO THE PPP FAMILY OF PHOSPHATASES.
 DR EMBL; Z77735; CAB01294.1; -;
 DR HSSP; P08129; 1FJM.
 DR InterPro; IPR004844; S/T_phosphatse.

DR ProDom; PD000252; S/T_phosphatse; 1.
 KW Hydrolase; Iron; Manganese.
 FT NON TER 1
 FT NON TER 24
 SQ SEQUENCE 24 AA; 2787 MW; F5418DAB76703E54 CRC64;

Query Match 61.3%; Score 19; DB 5; Length 24;
 Best Local Similarity 40.0%; Pred. No. 2.5e+02;
 Matches 4; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 YXXXXRLXER 10
 Db 2 YSDVRLFLDR 11

RESULT 4

Q63988 PRELIMINARY; PRT; 25 AA.
 AC Q63988;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Hox2.4 homeobox homolog protein (Fragment).
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OK NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TWISTAR;
 RC TISSUE=BONE MORPHOGENETIC PROTEIN-IMPLANTED SUBCUTANEOUS MUSCLE;
 RX MEDLINE=94271262; PubMed=7911662;
 RA Iimura T., Oida S., Takeda K., Maruoka Y., Sasaki S.;
 RT "Changes in homeobox-containing gene expression during ectopic bone
 formation induced by bone morphogenetic protein.";
 RL Biochem. Biophys. Res. Commun. 201:980-987(1994).
 DR EMBL; S71284; AAB31004.2; -;
 DR InterPro; IPR003356; Homeobox.
 DR Pfam; PF00046; homeobox; 1.
 KW DNA-binding; Homeobox; Nuclear protein.
 FT NON TER 1
 FT NON TER 25
 SQ SEQUENCE 25 AA; 3111 MW; 5B88C1D414ACB974 CRC64;
 Query Match 61.3%; Score 19; DB 11; Length 25;
 Best Local Similarity 80.0%; Pred. No. 2.6e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 RLXER 10
 Db 19 RLTER 23

RESULT 5
 Q9NKY0 PRELIMINARY; PRT; 38 AA.
 AC Q9NKY0;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Neoscox-1 homeodomain protein (Fragment).
 GN NEOSCOX-1.
 OS Aurelia aurita (moon jelly).
 OC Eukaryota; Metazoa; Cnidaria; Scyphozoa; Semeostomeae; Ulmaridae;
 OC Aurelia.
 OK NCBI_TaxID=6145;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Sakauchi M., Takahashi Y.;
 RT "Aurelia aurita neoscox-1 homeodomain protein, partial cds.";
 RL EMBL; AB037147; BAA89786.1; -;
 DR HSSP; P14653; 1B72.

DR InterPro: IPR001356; Homeobox.
 DR Pfam: PF00046; homeobox; 1.
 DR PRINTS: PR00024; HOMEBOX.
 DR ProDom: PD000010; Homeobox; 1.
 DR SMART: SM00389; Hox; 1.
 DR PROSITE: PS00071; HOMEBOX 2; 1.
 KW DNA-binding; Homeobox; Nuclear protein.
 FT NON_TER 1 1
 FT NON_TER 38 38
 SQ SEQUENCE 38 AA; 4904 MW; D78CC1FDCAE08B63 CRC64;

Query Match 61.3%; Score 19; DB 5; Length 38;
 Best Local Similarity 80.0%; Pred. No. 3.8e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 RLXER 10
 ||||
 Db 25 RLTER 29

RESULT 6

Q9VEZ8 PRELIMINARY; PRT; 52 AA.
 AC Q9VEZ8;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Pnr protein.
 GN PNR OR CG3978.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.D.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
 RA Ballwey R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Borchan M.R., Bouck J., Brockstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahike C., Davenport L.B., Davies P.,
 RA de Pablos K., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Flosser C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,

RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003711; AAF55264.2; -.
 DR FlyBase: FBgn0003117; pnr.
 DR InterPro: IPR001164; hRIP-like.
 DR InterPro: IPR000679; Znf_GATA.
 DR PRINTS: PR00619; GATAZNFINGER.
 DR PROSITE: PS00344; GATA_ZN_FINGER_1; 2.
 DR PROSITE: PS0114; GATA_ZN_FINGER_2; 2.
 SQ SEQUENCE 52 AA; 6046 MW; 6BA3775E6EE84346 CRC64;

Query Match 61.3%; Score 19; DB 5; Length 52;
 Best Local Similarity 80.0%; Pred. No. 5.1e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 RLXER 10
 ||||
 Db 22 RLSEK 26

RESULT 7

Q8UH70 PRELIMINARY; PRT; 53 AA.
 AC Q8UH70;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Hypothetical protein Atu0815.
 GN ATU0815 OR AGE_C1494.
 OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Rhizobium.
 OX NCBI_TaxID=176299;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21608550; PubMed=11743193;
 RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
 RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
 RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
 RA Chapman P., Clendinning J., Deatherage G., Gillet W., Grant C.,
 RA Kutayin T., Levy R., Li M.-J., McClelland E., Palmieri A., Gordon D.,
 RA Raymond C., Rouse G., Saenphimachak C., Wu Z., Romero P., Krespan W., Perry M.,
 RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
 RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
 RA Nester E.W.;
 RT "The genome of the natural genetic engineer Agrobacterium tumefaciens C58.";
 RL Science 294:2317-2323 (2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21608551; PubMed=11743194;
 RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M., Mullin L.,
 RA Qurollo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
 RA Hounell K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
 RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
 RA Flanagan C., Crowell C., Gursion J., Lomo C., Sear C., Strub G.,
 RA Cielo C., Slater S.;
 RT "Genome sequence of the plant pathogen and biotechnology agent Agrobacterium tumefaciens C58.";
 RL Science 294:2323-2328 (2001).
 DR EMBL: AE009048; AAL41829.1; -.
 DR EMBL: AE008014; AAK86622.1; -.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 53 AA; 6196 MW; DFB4E1C75C5B4CEP CRC64;

Query Match 61.3%; Score 19; DB 16; Length 53;
 Best Local Similarity 80.0%; Pred. No. 5.2e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 RLXER 10
 ||||

DB 24 RLAE 28

RESULT 8
027979 PRELIMINARY; PRT; 56 AA.

AC 027979; PRT; 56 AA.

DT 01-JAN-1998 (TRENBLREL. 05, Created)

DT 01-JAN-1998 (TRENBLREL. 05, Last sequence update)

DT 01-MAR-2002 (TRENBLREL. 20, Last annotation update)

DE Hypothetical protein AF2305.

GN AF2305.

OS Archaeoglobus fulgidus.

OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;

OC Archaeoglobaceae; Archaeoglobus.

OX NCBI_TaxID=2234;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;

RX MEDLINE=98049343; PubMed=9389475;

RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
Richardson D.L., Kervatage A.R., Graham D.E., Kyrleides N.C.,
Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Uterback T.,
Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,
Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
Maeson T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
Venter J.C.;

RA "The complete genome sequence of the hyperthermophilic, sulphate-
reducing archaeon Archaeoglobus fulgidus."

RL Nature 390:364-370 (1997).

DR EMBL; AE000945; AAB88955.1; -

DR TIGR; AF2305; -

KM Hypothetical protein; Complete proteome.

SQ SEQUENCE 56 AA; 6986 MW; DBDA20C268A7DP6 CRC64;

Query Match 61.3%; Score 19; DB 17; Length 56;
Best Local Similarity 80.0%; Pred. No. 5.5e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 6 RLXER 10
|||
Db 37 RLAE 41

RESULT 9
09Y188 PRELIMINARY; PRT; 58 AA.

AC 09Y188; PRT; 58 AA.

DT 01-NOV-1999 (TRENBLREL. 12, Created)

DT 01-NOV-1999 (TRENBLREL. 12, Last sequence update)

DT 01-MAR-2002 (TRENBLREL. 20, Last annotation update)

DE HBI homeodomain protein (Fragment).

GN HBI.

OS Priapulids caudatus.

OC Eukaryota; Metazoa; Priapulida; Priapulidae; Priapulidae; Priapulidae.

OX NCBI_TaxID=37621;

RN [1]

RP SEQUENCE FROM N.A.

RA MEDLINE=99318125; PubMed=10391241;
de Rosa R., Grenier J.K., Andreeva T., Cook C.E., Adoutte A., Akam M.,
Carroll S.B., Balavoine G.;

RA "Hox genes in brachiopods and priapulids and protostome evolution."

RL Nature 399:772-776 (1999).

CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).

DR EMBL; AF144888; AAD40644.1; -

DR HSSP; P02833; NANT.

DR InterPro; IPR001356; Homeobox.

DR InterPro; IPR000477; HTH_repressr.

DR Pfam; PF00046; Homeobox; 1.

DR PRINTS; PR00024; HOMEBOX.

DR PRINTS; PR00031; HTHREPRESSR.

DR ProDom; PD000010; Homeobox; 1.

DR SMART; SM00389; HOX; 1.

DR PROSITE; PS00027; HOMEBOX_1; 1.

DR PROSITE; PS00071; HOMEBOX_2; 1.

KW DNA-binding; Homeobox; Nuclear protein.

FT NON TER 1

SQ SEQUENCE 58 AA; 7323 MW; 572F30DA57C9A613 CRC64;

Query Match 61.3%; Score 19; DB 5; Length 58;
Best Local Similarity 80.0%; Pred. No. 5.7e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 6 RLXER 10
|||
Db 20 RLTER 24

RESULT 10
098M32 PRELIMINARY; PRT; 58 AA.

AC 098M32; PRT; 58 AA.

DT 01-OCT-2001 (TRENBLREL. 18, Created)

DT 01-OCT-2001 (TRENBLREL. 18, Last sequence update)

DT 01-MAR-2002 (TRENBLREL. 20, Last annotation update)

DE Hypothetical protein ms10755.

GN MS10755.

OS Rhizobium loti (Mesorhizobium loti).

OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;

OC Rhizobiaceae; Mesorhizobium.

OX NCBI_TaxID=381;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=MAFF303099;

RX MEDLINE=21082930; PubMed=11214968;

RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
Watanabe A., Idehara K., Ishikawa A., Kawahara K., Kimura T.,
Kishida Y., Kiyokawa C., Kohara M., Matsunoto M., Matsuno A.,
Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
Takeuchi C., Yamada M., Tabata S.;

RA "Complete genome structure of the nitrogen-fixing symbiotic bacterium
Mesorhizobium loti."

RL DNA Res. 7:331-338 (2000).

DR EMBL; AP002995; BAB48281.1; -

KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 58 AA; 6919 MW; C5BBDC3FF061623 CRC64;

Query Match 61.3%; Score 19; DB 16; Length 58;
Best Local Similarity 80.0%; Pred. No. 5.7e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 6 RLXER 10
|||
Db 12 RLTER 16

RESULT 11
092528 PRELIMINARY; PRT; 58 AA.

AC 092528; PRT; 58 AA.

DT 01-DEC-2001 (TRENBLREL. 19, Created)

DT 01-DEC-2001 (TRENBLREL. 19, Last sequence update)

DT 01-MAR-2002 (TRENBLREL. 20, Last annotation update)

DE Hypothetical protein lln1746.

GN Lln1746 OR Lln1251.

OS Listeria innocua.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;

OC Listeriaceae; Listeria.

OX NCBI_TaxID=1642;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CLIP 11262 / SEROVAR 6A;

RX PubMed=11679669;
 RA Glaser P., Frangeul L., Buchrieser C., Rueniok C., Amend A.,
 RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
 RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
 RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dusauguet O.,
 RA Entian K.-D., Faihi H., Garcia-del Portillo F., Garrido P.,
 RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
 RA Jones L.-M., Kaerst U., Krefit J., Kuhn M., Kunst F., Kurapkat G.,
 RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
 RA Nordieck G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
 RA Rammel B., Rose M., Schluter T., Simoes N., Tierrez A.,
 RA Vazquez-Boland J.-A., Voeg H., Wehland J., Cossart P.,
 RA "Comparative genomics of *Listeria species*.";
 RT Science 294:849-852(2001).
 DR EMBL; AL596169; CAC96977.1; -;
 DR EMBL; AL596168; CAC96482.1; -;
 DR Listitlist; LIN01251; -;
 DR Listitlist; LIN01746; -;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 58 AA; 6804 MW; 49412202B2B80E31 CRC64;

Query Match 61.3%; Score 19; DB 16; Length 58;
 Best Local Similarity 80.0%; Pred. No. 5.7e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 RLXER 10
 DB 11 RLSEER 15

RESULT 12

Q9BM19 ID Q9BM19 PRELIMINARY; PRT; 25 AA.
 AC Q9BM19;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE Gypsy-like reverse transcriptase (Fragment).
 OS Lissomyia mellita.
 OC Eukaryota; Metazoa; Echiura; Echiuroinea; Echiuridae; Thalamematinae;
 OC Lissomyia.
 OC NCBI_TaxID=145469;
 OX [1]
 RP SEQUENCE FROM N.A.
 RC TRANSPOSON=GRT-G1 RETROTRANSPOSON;
 RX MEDLINE=20570504; PubMed=11121049;
 RA Arkhipova I., Meselson M.;
 RT "Transposable elements in sexual and ancient asexual taxa.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:14473-14477(2000).
 DR EMBL; AY013983; AAG59959.1; -;
 KW RNA-directed DNA polymerase.
 FT NON_TER 1 1
 FT NON_TER 25 25
 SQ SEQUENCE 25 AA; 2856 MW; 14A04739D5E3AD5B CRC64;

Query Match 58.1%; Score 18; DB 5; Length 25;
 Best Local Similarity 80.0%; Pred. No. 4.9e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 RLXER 10
 DB 10 RLMEER 14

RESULT 13

Q9BM11 ID Q9BM11 PRELIMINARY; PRT; 26 AA.
 AC Q9BM11;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE Gypsy-like reverse transcriptase (Fragment).
 OS Spongillia lacustris (freshwater sponge).

OC Eukaryota; Metazoa; Porifera; Demospongiae; Ceractinomorpha;
 OC Haplosclerida; Spongillidae; Spongillia.
 OX NCBI_TaxID=6055;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TRANSPOSON=GRT-G4 RETROTRANSPOSON;
 RX MEDLINE=20570504; PubMed=11121049;
 RA Arkhipova I., Meselson M.;
 RT "Transposable elements in sexual and ancient asexual taxa.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:14473-14477(2000).
 DR EMBL; AY013993; AAG59967.1; -;
 KW RNA-directed DNA polymerase.
 FT NON_TER 1 1
 FT NON_TER 26 26
 SQ SEQUENCE 26 AA; 2956 MW; CDC8870829D053FB CRC64;

Query Match 58.1%; Score 18; DB 5; Length 26;
 Best Local Similarity 80.0%; Pred. No. 5e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 RLXER 10
 DB 10 RLMEER 14

RESULT 14

Q9BM76 ID Q9BM76 PRELIMINARY; PRT; 28 AA.
 AC Q9BM76;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE LINE-like reverse transcriptase (Fragment).
 OS Giardia lamblia (Giardia intestinalis).
 OC Eukaryota; Diplomonadida; Hexamitidae; Giardiae;
 OX NCBI_TaxID=5741;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TRANSPOSON=LRT-L3 RETROTRANSPOSON;
 RX MEDLINE=20570504; PubMed=11121049;
 RA Arkhipova I., Meselson M.;
 RT "Transposable elements in sexual and ancient asexual taxa.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:14473-14477(2000).
 DR EMBL; AY013917; AAG59902.1; -;
 KW RNA-directed DNA polymerase.
 FT NON_TER 1 1
 FT NON_TER 28 28
 SQ SEQUENCE 28 AA; 3304 MW; 6A3AF9C6F39667EF CRC64;

Query Match 58.1%; Score 18; DB 5; Length 28;
 Best Local Similarity 80.0%; Pred. No. 5.4e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 RLXER 10
 DB 18 RLSEER 22

RESULT 15

Q9BM75 ID Q9BM75 PRELIMINARY; PRT; 28 AA.
 AC Q9BM75;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE LINE-like reverse transcriptase (Fragment).
 OS Giardia lamblia (Giardia intestinalis).
 OC Eukaryota; Diplomonadida; Hexamitidae; Giardiae;
 OX NCBI_TaxID=5741;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TRANSPOSON=LRT-L4 RETROTRANSPOSON;
 RX MEDLINE=20570504; PubMed=11121049;

RA Arkhipova I., Meselson M.;
RT "Transposable elements in sexual and asexual taxa";
RL Proc. Natl. Acad. Sci. U.S.A. 97:1473-1477 (2000).
DR EMBL; AY013918; AAG59903.1; -
KW RNA-directed DNA polymerase.
FT NON_TER 1
FT NON_TER 28
SQ SEQUENCE 28 AA; 3269 MW; 6A305966533C67EF CRC64;

Query Match 58.1%; Score 18; DB 5; Length 28;
Best Local Similarity 80.0%; Pred. No. 5.4e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 6 REXER 10
Db 18 RLEER 22

Search completed: December 19, 2002, 16:21:52
Job time : 18.6471 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 19, 2002, 16:03:12 ; Search time 4.55882 Seconds
(without alignments)
90.980 Million cell updates/sec

Title: US-08-653-294C-39

Perfect score: 31

Sequence: 1 YXXXXRLXER 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 5116

Minimum DB seq length: 0

Maximum DB seq length: 60

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	58.1	37	1 24KD PLACH	P14592 plasmodium
2	18	58.1	47	1 AGRP_LUFUCY	P56568 luffa cylin
3	18	58.1	50	1 YORU_TTV1	P19294 thermoprote
4	16	51.6	28	1 FIBA CANFA	P02673 canis famil
5	16	51.6	45	1 VPU HV1N5	P08804 human immun
6	16	51.6	45	1 VPU HV1Z3	P08805 human immun
7	16	51.6	58	1 VPU HV1ZH	P08806 human immun
8	16	51.6	60	1 RS21_MYCPN	P57079 mycoplasma
9	16	51.6	60	1 Y569_ARCFU	O19199 archaeoglob
10	15	48.4	27	1 SECR_RABIT	P32647 oryctolagus
11	15	48.4	38	1 YF07_HAEIN	P44229 haemophilus
12	15	48.4	41	1 UL21_HSVE4	P24431 equine herp
13	15	48.4	54	1 SOX5_XENLA	P40647 xenopus lae
14	15	48.4	56	1 Y546_METJA	Q57966 methanococc
15	15	48.4	58	1 YBAS_PHAVU	P46380 amycolatops
16	14	45.2	20	1 AROQ_AMEYE	P30226 brassica na
17	14	45.2	23	1 AFP2_BRANA	P82976 mus musculu
18	14	45.2	25	1 GBX1_MOUSE	P30227 brassica ra
19	14	45.2	27	1 AFP1_BRARA	P30228 brassica ra
20	14	45.2	27	1 AFP2_BRARA	P56360 chlorella v
21	14	45.2	37	1 RK36_CHLVU	Q9muu8 mesostigma
22	14	45.2	38	1 RK36_MESVI	P20312 bacterioph
23	14	45.2	43	1 Y18_BPT3	P31901 alcaligenes
24	14	45.2	44	1 HYP1_ALCEU	O42367 brachydanio
25	14	45.2	45	1 HXB2_BRARE	P70936 bacillus me
26	14	45.2	45	1 TRPD_BACME	Q98r30 mycoplasma
27	14	45.2	46	1 R332_MYCPU	P80738 blepharisma
28	14	45.2	47	1 H4Y_BLEJA	Q44558 azotobacter
29	14	45.2	48	1 DPSD_AZOVI	P31256 xenopus lae
30	14	45.2	48	1 HXB2_XENLA	O28197 archaeoglob
31	14	45.2	48	1 YK82_ARCFU	Q28599 ovis aries
32	14	45.2	49	1 HXA5_SHEEP	P35870 bacillus li
33	14	45.2	49	1 R332_BACLI	

34 14 45.2 51 1 APP1_SINAL P30231 sinapis alb
35 14 45.2 54 1 SX13_XENLA P40650 xenopus lae
36 14 45.2 54 1 YHAI_ECOLI P42625 escherichia
37 14 45.2 55 1 PHNS_DESVH Q06173 desulfovibr
38 14 45.2 55 1 SX19_MOUSE Q62249 mus musculu
39 14 45.2 56 1 Y6KD_STRLI P22400 streptomyce
40 14 45.2 58 1 CCKN_BOVIN P41520 bos taurus
41 14 45.2 58 1 COX9_YEAST P07255 saccharomyc
42 14 45.2 58 1 IHH_DANAU P79693 carassius a
43 14 45.2 58 1 IHH_DANAT O13240 danio aff.
44 14 45.2 58 1 IHH_DANKE P79711 danio kerri
45 14 45.2 58 1 IHH_DANPU P79719 danio pulch

ALIGNMENTS

RESULT 1
24KD_PLACH STANDARD; PRT; 37 AA.
ID AC P14592;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 01-MAY-1991 (Rel. 18, Last annotation update)
DE 24 Kda antigen (Fragment).
OS Plasmodium chabaudi.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5825;
RN [1]
RX MEDLINE=87174746; PubMed=3550696;
RA Langley G., Sibilli L., Mattei D., Falanga P., Mercereau-Puijalon O.;
RT "karyotype comparison between P. chabaudi and P. falciparum: analysis
of a P. chabaudi cDNA containing sequences highly repetitive in P.
falciparum";
RL Nucleic Acids Res. 15:2203-2211 (1987).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; X04892; CAA28580.1; -.
KW Malaria.
FT NON_TER 1
SQ SEQUENCE 37 AA; 4757 MW; E6F0EB6A45CFEBD5 CRC64;
Query Match 58.1%; Score 18; DB 1; Length 37;
Best Local Similarity 44.4%; Pred. No. 78;
Matches 4; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 YXXXXRLXEX 9
DB 21 YTVAQRLLE 29
RESULT 2
AGRP_LUFUCY STANDARD; PRT; 47 AA.
ID AGRP_LUFUCY
AC P56568;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE 6.5 kDa arginine/glutamate-rich polypeptide (6.5K-AGRP).
OS Luffa cylindrica (Smooth loofah) (Sponge gourd).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; Rosidae;
OC Eurosoidae I; Cucurbitales; Cucurbitaceae; Luffa.
OX NCBI_TaxID=3670;
RN [1]

```

RP SEQUENCE.
RC TISSUE=Seed;
RA MEDLINE=97357433; PubMed=9214759;
RA Kimura M., Park S.S., Sakai R., Yamasaki N., Funatsu G.,
RT "Primary structure of 6.5k-arginine/glutamate-rich polypeptide from
RT the seeds of sponge gourd (Luffa cylindrica).";
RL Biosci. Biotechnol. Biochem. 61:984-988(1997).
CC -1- FUNCTION: STORAGE PROTEIN WHICH PROVIDES NITROGEN AND CARBON
CC RESERVES DURING GERMINATION AND SEEDLING GROWTH.
CC -1- MASS SPECTROMETRY: MW=5693.39; METHOD=MALDI.
CC -1- SIMILARITY: SOME, TO 7S SEED STORAGE PROTEINS.
KM Seed storage protein.
FT DISULFID 12 33
FT DISULFID 16 29
SQ SEQUENCE 47 AA; 5698 MW; 588B0EC62273AC05 CRC64;

Query Match 58.1%; Score 18; DB 1; Length 47;
Best Local Similarity 80.0%; Pred. No. 1e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 6 RLXER 10
DB 36 RLXER 40

RESULT 3
YORU_TTV1 STANDARD; PRT; 50 AA.
ID_YORU_TTV1
AC 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DE 01-NOV-1990 (Rel. 16, Last annotation update)
DE Hypothetical 6.1 kDa protein.
OS Thermoproteus tenax virus 1 (strain KRA1) (TTV1).
CC Viruses; dsDNA viruses, no RNA stage; Lipothirixviridae;
CC Lipothirixvirus.
CC NCBI_TaxID=10480;
KM (1)
RP SEQUENCE FROM N.A.
RA Neumann H.;
RL Submitted (MAR-1989) to the EMBL/Genbank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X14855; CAA32990.1; -
KM Hypothetical protein.
SQ SEQUENCE 50 AA; 6115 MW; 4308C8E0937B4DCF CRC64;

Query Match 58.1%; Score 18; DB 1; Length 50;
Best Local Similarity 80.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 6 RLXER 10
DB 18 RLXER 22

RESULT 4
FTBA_CANFA STANDARD; PRT; 28 AA.
ID_FTBA_CANFA
AC P02673; P14464;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Fibrinogen alpha chain (Contains: Fibrinopeptide A) (Fragment).
GN FGA.
OS Canis familiaris (Dog), and

```

```

OS Vulpes vulpes (Red fox).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615, 9627;
RN (1)
RP SEQUENCE.
RC SPECIES=C.familiaris;
RA MEDLINE=76081726; PubMed=1198547;
RA Birken S., Wilner G.D., Canfield R.E.;
RT "Studies of the structure of canine fibrinogen.";
RL Thromb. Res. 7:599-610(1975).
RN (2)
RP SEQUENCE OF 1-16.
RC SPECIES=C.familiaris, and V.vulpes;
RA Blomback B., Blomback M., Grondahl N.J.;
RT "Studies on fibrinopeptides from mammals.";
RL Acta Chem. Scand. 19:1789-1791(1965).
RN (3)
RP SEQUENCE OF 1-16.
RC SPECIES=C.familiaris;
RA Osbahr A.J. Jr., Colman R.W., Laki K., Gladner J.A.;
RT "The nature of the peptides released from canine fibrinogen.";
RL Biochem. Biophys. Res. Commun. 14:555-558(1964).
CC -1- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT
CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET
CC AGGREGATION.
CC -1- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
CC -1- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.
DR PIR; A03118; A03118.
DR PIR; A05296; A05296.
KM Blood coagulation; Plasma; Phosphorylation.
FT PEPTIDE 1 16 FIBRINOPEPTIDE A.
FT MOD_RES 3 3 PHOSPHORYLATION (IN SOME MOLECULES).
FT CONFLICT 2 2 N->D (IN REF. 2).
FT CONFLICT 4 7 KEKG -> EKGQ (IN REF. 2).
FT NON_TER 28
SQ SEQUENCE 28 AA; 2958 MW; 09DCD3F923BFEBD2 CRC64;

Query Match 51.6%; Score 16; DB 1; Length 28;
Best Local Similarity 60.0%; Pred. No. 2.2e+02;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 6 RLXER 10
DB 19 RIVER 23

RESULT 5
VPU_HVINS STANDARD; PRT; 45 AA.
ID_VPU_HVINS
AC P08804;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE VPU protein (U ORF protein) (Fragment).
GN VPU.
OS Human immunodeficiency virus type 1 (New York-5 isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11698;
RN (1)
RP SEQUENCE FROM N.A.
RA MEDLINE=86259728; PubMed=3014529;
RA Willey R.W., Rutledge R.A., Dias S., Folks T., Theodore T.,
RA Buckler C.E., Martin M.A.;
RT "Identification of conserved and divergent domains within the
RT envelope gene of the acquired immunodeficiency syndrome retrovirus.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:5038-5042(1986).
CC -1- FUNCTION: ACTS IN THE DEGRADATION OF CD4 IN THE ENDOPLASMIC

```



```
CC RETICULUM AND IN THE ENHANCEMENT OF VIRION RELEASE FROM THE PLASMA
CC MEMBRANE OF INFECTED CELLS.
CC -1- SUBCELLULAR LOCATION: Membrane-bound.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: K03346; AAB02406.1; -.
CC DR HSSP: P19554; 1VPU.
CC DR HIV: K03346; VFUS23.
CC DR InterPro: IPR002094; Vpu.
CC DR Pfam: PF00558; Vpu; 1.
CC KW Transmembrane; AIDS.
CC FT NON_TER 1 1
CC SEQUENCE 45 AA; 5166 MW; 2AA46D116BB624F7 CRC64;
CC
CC Query Match 51.6%; Score 16; DB 1; Length 45;
CC Best Local Similarity 60.0%; Pred. No. 3.6e+02;
CC Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
CC
CC QY 6 RLXER 10
CC Db 9 RIRER 13
CC
CC RESULT 6
CC VPU_HV123 STANDARD; PRT; 45 AA.
CC ID VPU_HV123
CC AC P08005;
CC DT 01-NOV-1988 (Rel. 09, Created)
CC DT 01-NOV-1988 (Rel. 09, Last sequence update)
CC DT 01-NOV-1997 (Rel. 35, Last annotation update)
CC DE VPU protein (U ORF protein) (Fragment).
CC GN VPU.
CC OS Human immunodeficiency virus type 1 (Zaire 3 isolate) (HIV-1).
CC OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
CC OX NCBI_TaxID=11680;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RX MEDLINE=86259728; PubMed=3014529;
CC RA Willey R.W., Rutledge R.A., Dias S., Folks T., Theodore T.,
CC RA Buckler C.E., Martin M.A.;
CC RT "Identification of conserved and divergent domains within the
CC RT envelope gene of the acquired immunodeficiency syndrome retrovirus.";
CC RL Proc. Natl. Acad. Sci. U.S.A. 83:5038-5042(1986).
CC CC -1- FUNCTION: ACTS IN THE DEGRADATION OF CD4 IN THE ENDOPLASMIC
CC CC RETICULUM AND IN THE ENHANCEMENT OF VIRION RELEASE FROM THE PLASMA
CC CC MEMBRANE OF INFECTED CELLS.
CC CC -1- SUBCELLULAR LOCATION: Membrane-bound.
CC CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: K03347; AAA45372.1; -.
CC DR HSSP: P19554; 1VPU.
CC DR HIV: K03347; VFUS23.
CC DR InterPro: IPR002094; Vpu.
CC DR Pfam: PF00558; Vpu; 1.
CC KW Transmembrane; AIDS.
CC FT NON_TER 1 1
CC SEQUENCE 45 AA; 5342 MW; E9FF1EEAA174FA49 CRC64;
CC
CC Query Match 51.6%; Score 16; DB 1; Length 45;
```

```
Best Local Similarity 60.0%; Pred. No. 3.6e+02;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 6 RLXER 10
Db 9 RIRER 13

RESULT 7
VPU_HV12H STANDARD; PRT; 58 AA.
ID VPU_HV12H
AC P08006;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE VPU protein (U ORF protein) (Fragment).
GN VPU.
OS Human immunodeficiency virus type 1 (Zaire HZ321 isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11692;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89228766; PubMed=2713163;
RA Srinivasan A., York D., Butler D. Jr., Jannoun-Nasr R., Getchell J.,
RA McCormick J., Ou C.Y., Myers G., Smith T., Chen E.;
RT "Molecular characterization of HIV-1 isolated from a serum collected
RT in 1976; nucleotide sequence comparison to recent isolates and
RT generation of hybrid HIV.";
RL AIDS Res. Hum. Retroviruses 5:121-129(1989).
CC -1- FUNCTION: ACTS IN THE DEGRADATION OF CD4 IN THE ENDOPLASMIC
CC CC RETICULUM AND IN THE ENHANCEMENT OF VIRION RELEASE FROM THE PLASMA
CC CC MEMBRANE OF INFECTED CELLS.
CC CC -1- SUBCELLULAR LOCATION: Membrane-bound.
CC CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M15896; AAB53947.1; -.
CC DR HSSP: P19554; 1VPU.
CC DR HIV: M15896; VFUS2321.
CC DR InterPro: IPR002094; Vpu.
CC DR Pfam: PF00558; Vpu; 1.
CC KW Transmembrane; AIDS.
CC FT NON_TER 1 1
CC SEQUENCE 58 AA; 6789 MW; 40EE98A77BE3DACE CRC64;
CC
CC Query Match 51.6%; Score 16; DB 1; Length 58;
CC Best Local Similarity 60.0%; Pred. No. 4.7e+02;
CC Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 6 RLXER 10
Db 22 RIRER 26

RESULT 8
RS21_MYCPN STANDARD; PRT; 60 AA.
ID RS21_MYCPN
AC P57079;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 30S ribosomal protein S21.
GN RPSU OR MPN296 OR MP539.1.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=21104;
RN [1]
```

```

RP  SEQUENCE FROM N.A.
RC  STRAIN=ATCC 29342 / M129;
RA  MEDLINE=97105883; PubMed=8948633;
RA  Himmelfarb R., Hilbert H., Plagens H., Pirkl E., Li B.-C.,
RA  Herrmann R.;
RT  "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT  pneumoniae.";
RL  Nucleic Acids Res. 24:4420-4449(1996).
RN  [2]
RP  IDENTIFICATION.
RX  MEDLINE=20411492; PubMed=10954595;
RA  Dandekar T., Huynen M., Regula J.T., Ueberle B., Zimmermann C.U.,
RA  Andrade M.A., Doerks T., Sanchez-Pulido L., Snel B., Suyama M.,
RA  Yuan Y.P., Herrmann R., Bork P.;
RT  "Re-annotating the Mycoplasma pneumoniae genome sequence: adding
RT  value, function and reading frames.";
RL  Nucleic Acids Res. 28:3278-3288(2000).
CC  -1 SIMILARITY: BELONGS TO THE S21P FAMILY OF RIBOSOMAL PROTEINS.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
CC  EMBL; AE000052; -; NOT ANNOTATED_CDS.
DR  EMBL; AE000053; AAG34754.1; -
DR  InterPro; IPR001911; Ribosomal_S21.
DR  PROSITE; PS01181; RIBOSOMAL_S21; FALSE_NEG.
KW  Ribosomal protein; Complete proteome.
SQ  SEQUENCE 60 AA; 7515 MW; 30D87DF4635F003 CRC64;

Query Match          51.6%; Score 16; DB 1; Length 60;
Best Local Similarity 60.0%; Pred. No. 4.8e+02;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY  6 RLXER 10
   |||
   27 RLXER 31

DB  27 RLXER 31

RESULT 9
ID  Y569 ARCFU STANDARD; PRT; 60 AA.
AC  O19159;
DT  30-MAY-2000 (Rel. 39, Created)
DT  30-MAY-2000 (Rel. 39, Last sequence update)
DT  16-OCT-2001 (Rel. 40, Last annotation update)
DE  Hypothetical protein AF0569.
GN  AF0569.
OS  Archaeoglobus fulgidus.
OC  Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC  Archaeoglobaceae; Archaeoglobus.
OX  NCBI_TaxID=2234;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX  MEDLINE=98049343; PubMed=9389475;
RA  Kleink H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA  Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA  Richardson D.L., Kierulff A.R., Graham D.E., Kyriakides N.C.,
RA  Fleischman R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA  Kirschner E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA  Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA  Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Uterback T.,
RA  Cotton P.M.D., Spriggs T., Artiach P., Kalne B.P., Sykes S.M.,
RA  Sadov P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA  Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA  Venter J.C.;
RT  "The complete genome sequence of the hyperthermophilic, sulphate-
RT  reducing archaeon Archaeoglobus fulgidus.";
```

```

RL  Nature 390:364-370(1997).
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
CC  EMBL; AE001065; AAB90673.1; -
DR  TIGR; AF0569; -
KW  Hypothetical protein; Complete proteome.
SQ  SEQUENCE 60 AA; 7054 MW; B62833AD3A00788B CRC64;

Query Match          51.6%; Score 16; DB 1; Length 60;
Best Local Similarity 60.0%; Pred. No. 4.8e+02;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY  6 RLXER 10
   |||
   10 RLXER 14

DB  10 RLXER 14

RESULT 10
ID  SECR_RABIT STANDARD; PRT; 27 AA.
AC  P32647;
DT  01-OCT-1993 (Rel. 27, Created)
DT  01-OCT-1993 (Rel. 27, Last sequence update)
DT  01-NOV-1995 (Rel. 32, Last annotation update)
DE  Secretin.
GN  SCT.
OS  Oryctolagus cuniculus (Rabbit).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX  NCBI_TaxID=9986;
RN  [1]
RP  SEQUENCE.
RP  TISSUE=Small intestine;
RX  MEDLINE=90259845; PubMed=234298;
RX  Goosen D., Buscail L., Cauvin A., Goulet P., de Neef P., Rathe J.,
RA  Robberecht P., Vandermere-Piret M.C., Vandemeers A., Christophe J.;
RT  "Amino acid sequence of VIP, PHI and secretin from the rabbit small
RT  intestine.";
RL  Peptides 11:123-128(1990).
CC  -1 FUNCTION: STIMULATES FORMATION OF NAHCO(3)-RICH BILE AND INHIBITS HCL PRODUCTION
CC  AND SECRETION OF NAHCO(3)-RICH BILE AND INHIBITS HCL PRODUCTION
CC  BY THE STOMACH.
CC  -1 SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
CC  PIR; C60415; C60415.
DR  InterPro; IPR000532; Glucagon.
DR  Pfam; PF00123; hormone2; 1.
DR  PRINTS; PR00275; GLUCAGON.
DR  SMART; SM00070; GLUCA; 1.
DR  PROSITE; PS00260; GLUCAGON; 1.
KW  Glucagon family; Hormone; Amidation.
FT  MOD RES 27
FT  MOD RES 27
SQ  SEQUENCE 27 AA; 3105 MW; 38A01580BDD3618 CRC64;

Query Match          48.4%; Score 15; DB 1; Length 27;
Best Local Similarity 60.0%; Pred. No. 4.1e+02;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY  6 RLXER 10
   |||
   12 RLXER 16

DB  12 RLXER 16

RESULT 11
ID  YF07_HAEIN STANDARD; PRT; 38 AA.
AC  P44229;
```

DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Very hypothetical protein HI1507 in Mu-like prophage FluMu region.
 GN HI1507.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC Haemophilus.
 OX NCBI_TaxID=727;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Rd / KW20 / ATCC 51907;
 RX MEDLINE=95350630; PubMed=7542800;
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 RA Utterback T.R., Hanna M.C., Spriggs T., Saudek D.M., Brandon R.C.,
 RA Pine L.D., Fritchman J.L., Fuhrmann J.B., Saudek D.M., Brandon R.C.,
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.;
 RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
 RT Rd.";
 RL Science 269:496-512(1995).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U32827; AAC23166.1; -
 DR TIGR; HI1507; -
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 38 AA; 4123 MW; C3DE70D18D0BE11 CRC64;
 Query Match 48.4%; Score 15; DB 1; Length 38;
 Best Local Similarity 60.0%; Pred. No. 5.8e+02;
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 6 RLXER 10
 :| ||
 Db 7 QLTER 11

 RESULT 12
 ID UL21 HSV4 STANDARD; PRT; 41 AA.
 AC P24431;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 01-APR-1993 (Rel. 25, Last annotation update)
 DE Protein UL21 homolog (Fragment).
 OS Equine herpesvirus type 4 (strain 1942) (EHV-4) (Equine herpesvirus
 OS type 1 subtype 2).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicellovirus.
 OX NCBI_TaxID=10333;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=90362066; PubMed=2167933;
 RA Nicolson L., Cullinan A.A., Onions D.E.;
 RT "The nucleotide sequence of an equine herpesvirus 4 gene homologue of
 RT the herpes simplex virus 1 glycoprotein H gene.";
 RL J. Gen. Virol. 71:1793-1800(1990).
 CC -!- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL21,
 CC EHV-1 40, EHV-4 UL21, AND VZV 38.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D14486; BAA03380.1; -
 DR InterPro; IPR004936; Herpes_UL21.
 DR Pfam; PF03252; UL21; 1.
 FT NON_TER 1
 SQ SEQUENCE 41 AA; 4683 MW; 341D6EE97859349F CRC64;
 Query Match 48.4%; Score 15; DB 1; Length 41;
 Best Local Similarity 40.0%; Pred. No. 6.3e+02;
 Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 1 YXXXXLXER 10
 :| ||
 Db 19 YLYYLLSER 28

 RESULT 13
 ID SOX5 XENLA STANDARD; PRT; 54 AA.
 AC P40847;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Transcription factor SOX-5 (Fragment).
 GN SOX-5.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92310993; PubMed=1614875;
 RA Denny P., Swift S., Brand N., Dabhadre N., Barton P., Ashworth A.;
 RT "A conserved family of genes related to the testis determining gene,
 RT SRY.";
 RL Nucleic Acids Res. 20:2887-2887(1992).
 CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
 CC -!- SIMILARITY: CONTAINS 1 HMG BOX.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X65653; CAA46604.1; -
 DR PIR; S21492; S21492.
 DR PIR; S22949; S22949.
 DR HSRP; Q05066; 1HRX.
 DR InterPro; IPR000910; HMG 12_box.
 DR Pfam; PF00505; HMG_box; 1.
 DR SMART; SM00398; HMG; 1.
 KW DNA-binding; Nuclear protein; Transcription regulation.
 FT NON_TER 1
 FT DNA_BIND <1 >54 HMG BOX.
 FT NON_TER 54
 SQ SEQUENCE 54 AA; 6536 MW; E08A1BFC00B3EDD8 CRC64;
 Query Match 48.4%; Score 15; DB 1; Length 54;
 Best Local Similarity 30.0%; Pred. No. 8.4e+02;
 Matches 3; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
 QY 1 YXXXXLXER 10
 :| ||
 Db 40 YEEQARLSQ 49

```

RESULT 14
Y546_METUA STANDARD; PRT; 56 AA.
ID Y546_METUA
AC Q57966;
DT 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Hypothetical protein MJ0546.
GN MJ0546.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OK NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kervatage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weisscock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Meldrum J.F., Fuhrmann J.L., Nguyen D.,
RA Uterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Corton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii."
RL Science 273:1058-1073 (1996).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U67504; AAB98548.1; -.
DR TIGR; MJ0546; -.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 2 22 POTENTIAL.
FT DOMAIN 27 47 GUU-RICH.
SQ SEQUENCE 56 AA; 6460 MW; 66CBFC395548BA5D CRC64;

Query Match 48.4%; Score 15; DB 1; Length 56;
Best Local Similarity 60.0%; Pred. No. 8.7e+02;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 6 RLXR 10
: |||
Db 44 KLINR 48

RESULT 15
YBAS_PHAUV STANDARD; PRT; 58 AA.
ID YBAS_PHAUV
AC P03939;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 21-JUL-1986 (Rel. 01, Last annotation update)
DE Hypothetical basic polypeptide.
OS Phaseolus vulgaris (Kidney bean) (French bean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Phaseolus.
OK NCBI_TaxID=3885;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Tendergreen;
RX MEDLINE=85008540; PubMed=6090563;
RA Hoffman L.M.;

```

```

RT "Structure of a chromosomal Phaseolus vulgaris lectin gene and its
RT transcript."
RL J. Mol. Appl. Genet. 2:447-453 (1984).
DR PIR; A04525; Q0FB.
KW Hypothetical protein.
SQ SEQUENCE 58 AA; 6859 MW; DA6BF0F01DB67D5E CRC64;

Query Match 48.4%; Score 15; DB 1; Length 58;
Best Local Similarity 60.0%; Pred. No. 9e+02;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 6 RLXR 10
: |||
Db 48 RLIR 52

Search completed: December 19, 2002, 16:19:44
Job time : 5.55882 secs

```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 19, 2002, 16:16:02 ; Search time 8.67647 Seconds
(without alignments)
110.799 Million cell updates/sec

Title: US-08-653-294C-39
Perfect score: 31
Sequence: 1 YXXXXRLXER 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 15911

Minimum DB seq length: 0
Maximum DB seq length: 60

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 73:*
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	23	74.2	56	2 B89903	hypothetical prote
2	19	61.3	24	2 T42257	phosphoprotein pho
3	19	61.3	53	2 E97458	hypothetical prote
4	19	61.3	53	2 AG2676	hypothetical prote
5	19	61.3	56	2 A69538	conserved hypothet
6	19	61.3	58	2 AB1589	hypothetical prote
7	19	61.3	58	2 AI1650	hypothetical prote
8	18	58.1	47	2 JCS557	arginine/glutamate
9	18	58.1	48	2 A84247	hypothetical prote
10	18	58.1	48	2 T01745	hypothetical prote
11	18	58.1	48	2 T01702	hypothetical prote
12	18	58.1	58	2 S66328	protein kinase AK1
13	18	58.1	58	2 E84008	hypothetical prote
14	17	54.8	24	2 T42259	phosphoprotein pho
15	17	54.8	39	2 C87422	hypothetical prote
16	17	54.8	39	2 G97623	hypothetical prote
17	17	54.8	42	2 A61600	segmentation prote
18	17	54.8	52	2 T00166	hypothetical prote
19	17	54.8	52	2 S66323	protein kinase AK1
20	17	54.8	58	2 S77801	probable 6-phospho
21	16	51.6	28	2 A05296	fibronogen alpha c
22	16	51.6	30	2 S25666	phosphopyruvate hy
23	16	51.6	43	2 S23641	cytochrome-b5 redu
24	16	51.6	48	2 A83760	hypothetical prote
25	16	51.6	52	4 S70252	hypothetical prote
26	16	51.6	53	2 S00391	hypothetical prote
27	16	51.6	55	2 A83286	hypothetical prote
28	16	51.6	56	2 S66324	protein kinase AK1
29	16	51.6	57	2 D90828	hypothetical prote

30	16	51.6	59	2 D82505	hypothetical prote
31	16	51.6	59	2 T12913	hypothetical prote
32	16	51.6	60	2 A69321	DR-beta chain MHC
33	16	51.6	60	2 D85758	hypothetical prote
34	15	48.4	12	2 S49547	hypothetical prote
35	15	48.4	13	2 PC2371	probable endopepti
36	15	48.4	15	2 A56786	pimeloyl-CoA synth
37	15	48.4	20	2 S50741	probable trypsin i
38	15	48.4	21	2 S78575	protein kinase C i
39	15	48.4	21	2 S78574	protein kinase C i
40	15	48.4	23	2 S60564	homeodomain protei
41	15	48.4	26	2 S62672	transcobalamin - b
42	15	48.4	27	2 C60415	secretin - rabbit
43	15	48.4	27	2 I49747	phycobilisome beta
44	15	48.4	30	2 F60691	hypothetical prote
45	15	48.4	35	2 F69827	hypothetical prote

ALIGNMENTS

RESULT 1

B89903
hypothetical protein [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: B89903
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogu
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: B89903
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-56 <XUR>
A:Cross-references: GB:BA000018; PID:g13701087; PIDN:BAB42382.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: SAS037

Query Match 74.2%; Score 23; DB 2; Length 56;
Best Local Similarity 50.0%; Pred. No. 9;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 YXXXXRLXER 10
Db 21 YINAYRLAER 30

RESULT 2

T42257
phosphoprotein phosphatase (EC 3.1.3.16) - Caenorhabditis elegans (fragment)
C:Species: Caenorhabditis elegans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 02-Jun-2000
C:Accession: T42257
R:Zeke, T.; Gergely, P.; Dombradi, V.
submitted to the EMBL Data Library, July 1996
A:Description: The catalytic subunits of Ser/Thr protein phosphatases from Caenorhabdit
A:Reference number: Z22131
A:Accession: T42257
A>Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-24 <ZEK>
A:Cross-references: EMBL:Z77735; PIDN:CAB01294.1
C:Superfamily: phosphoprotein phosphatase; phosphoesterase core homology; phosphoprotei
C:Keywords: phosphoric monoester hydrolase

Query Match 61.3%; Score 19; DB 2; Length 24;
Best Local Similarity 40.0%; Pred. No. 57;
Matches 4; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 1 YXXXXRLXER 10
Db 2 YSDVRLRPFDR 11

RESULT 3
E97458
[hypothetical protein AGR_C_1494 [imported] - Agrobacterium tumefaciens (strain C58, Cere
C/Species: Agrobacterium tumefaciens
C/Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 11-Jan-2002
C/Accession: E97458
R/Goodner, B.; Hinkle, G.; Galling, S.; Miller, N.; Blanchard, M.; Ourullo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A/Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A/Reference number: A97359; PMID:11743194
A/Accession: E97458
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-53 <KUR>
A/Cross-references: GB:AE007869; PIDN:AAK6622.1; PID:g15155796; GSPDB:GN00169
C/Genetics:
A/Gene: AGR_C_1494
A/Map position: circular chromosome

Query Match 61.3%; Score 19; DB 2; Length 53;
Best Local Similarity 80.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 6 RLXER 10
Db 24 RLAE 28

RESULT 4
AG2676
[hypothetical protein Atu0815 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C/Species: Agrobacterium tumefaciens
C/Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
C/Accession: AG2676
R/wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kuyavlin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A/Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A/Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A/Reference number: AB2577; PMID:11743193
A/Accession: AG2676
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-53 <KUR>
A/Cross-references: GB:AE006888; PIDN:ALA1829.1; PID:g17739187; GSPDB:GN00186
C/Genetics:
A/Experimental source: strain C58 (Dupont)
A/Gene: Atu0815
A/Map position: circular chromosome

Query Match 61.3%; Score 19; DB 2; Length 53;
Best Local Similarity 80.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 6 RLXER 10
Db 24 RLAE 28

RESULT 5
A69538
[conserved hypothetical protein AF2305 - Archaeoglobus fulgidus
C/Species: Archaeoglobus fulgidus
C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
C/Accession: A69538

R/Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A/Authors: Uitterlinden, T.; Cotton, M.D.; Spriggs, T.; Artach, P.; Kaine, B.P.; Sykes, S.J.
Smith, H.O.; Woese, C.R.; Venturi, J.C.
A/Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
A/Reference number: A69250; MUID:96049343; PMID:9389475
A/Accession: A69538
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-56 <KLE>
A/Cross-references: GB:AE000945; GB:AE000782; NID:92669268; PIDN:AB88955.1; PID:g264821

Query Match 61.3%; Score 19; DB 2; Length 56;
Best Local Similarity 80.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 6 RLXER 10
Db 37 RLAE 41

RESULT 6
AB1589
[hypothetical protein lin1251 [imported] - Listeria innocua (strain Clp11262)
C/Species: Listeria innocua
C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C/Accession: AB1589
R/Glaser, P.; Prangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurge, O.; Entian, K.D.; Fsihl, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A/Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapk, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schluter, T.; Simoes, N.; Tietze, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend,
A/Title: Comparative genomics of Listeria species.
A/Reference number: AB1077; MUID:21537279; PMID:11679669
A/Accession: AB1589
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-58 <GLA>
A/Cross-references: GB:AL592022; PIDN:CAC96482.1; PID:g16413725; GSPDB:GN00178
A/Experimental source: strain Clp11262
C/Genetics:
A/Gene: lin1251

Query Match 61.3%; Score 19; DB 2; Length 58;
Best Local Similarity 80.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 6 RLXER 10
Db 11 RLSE 15

RESULT 7
A11650
[hypothetical protein lin1746 [imported] - Listeria innocua (strain Clp11262)
C/Species: Listeria innocua
C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C/Accession: A11650
R/Glaser, P.; Prangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurge, O.; Entian, K.D.; Fsihl, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A/Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapk, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schluter, T.; Simoes, N.; Tietze, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend,
A/Title: Comparative genomics of Listeria species.
A/Reference number: AB1077; MUID:21537279; PMID:11679669
A/Accession: A11650
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-58 <GLA>

A;Cross-references: GB:AL592022; PIDN:CAC96977.1; PID:g16414233; GSPDB:GN00178
A;Experimental source: strain Clp11262
C;Genetics:
A;Gene: lin1746

Query Match 61.3%; Score 19; DB 2; Length 58;
Best Local Similarity 80.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 RLXER 10
|||
Db 11 RLSEER 15

RESULT 8

JC5557
arginine/glutamate-rich 6.5K polypeptide - smooth loofah

C;Species: Luffa cylindrica (smooth loofah)
C;Date: 02-Sep-1997 #sequence_revision 05-Sep-1997 #text_change 07-May-1999

C;Accession: JC5557
R;Kimura, M.; Park, S.S.; Sakai, R.; Yamasaki, N.; Funatsu, G.

A;Title: Primary structure of 6.5K-arginine/glutamate-rich polypeptide from the seeds of
A;Reference number: JC5557; MUID:97357433; PMID:9214759
A;Accession: JC5557

A;Molecule type: protein

A;Residues: 1-47 <KIM>

A;Experimental source: seed

C;Comment: This protein is a storage protein which provides nitrogen and carbon reserves
F;12-33,16-29/Disulfide bonds: #status predicted

Query Match 58.1%; Score 18; DB 2; Length 47;
Best Local Similarity 80.0%; Pred. No. 2.3e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 RLXER 10
|||
Db 36 RLREER 40

RESULT 9

A84247
hypothetical protein Vng0907h [imported] - Halobacterium sp. NRC-1

C;Species: Halobacterium sp. NRC-1

C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001

C;Accession: A84247

R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.
; Leithauer, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo
Jung, K.H.; Alam, M.; Freitas, T.

Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000

A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A;Title: Genome sequence of Halobacterium species NRC-1.

A;Reference number: A84160; MUID:20504483; PMID:11016950

A;Accession: A84247

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-48 <STO>

A;Cross-references: GB:AE004437; NID:g10580467; PIDN:AAG19341.1; GSPDB:GN00138

C;Genetics:

A;Gene: VNG0907H

Query Match 58.1%; Score 18; DB 2; Length 48;
Best Local Similarity 80.0%; Pred. No. 2.3e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 RLXER 10
|||
Db 35 RLREER 39

RESULT 10

T01745

hypothetical protein 48 - maize mitochondrion

C;Species: mitochondrion Zea mays (maize)
C;Date: 19-Feb-1999 #sequence_revision 19-Feb-1999 #text_change 29-Oct-1999
C;Accession: T01745
R;Zheng, D.; Daniell, H.
submitted to the EMBL Data Library, December 1994
A;Description: Identification of tRNA^{Leu} (CAT), orf31 and orf48 of chloroplast origin in
A;Reference number: Z14415
A;Accession: T01745
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-48 <ZHE>
A;Cross-references: EMBL:U18301; NID:g664761; PIDN:AAA83741.1; PID:g664762
C;Genetics:
A;Gene: ORF48
A;Genome: mitochondrion
C;Keywords: mitochondrion

Query Match 58.1%; Score 18; DB 2; Length 48;
Best Local Similarity 80.0%; Pred. No. 2.3e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 RLXER 10
|||
Db 36 RLLEER 40

RESULT 11

T01702

hypothetical protein 48 - maize mitochondrion (fragment)

C;Species: mitochondrion Zea mays (maize)
C;Date: 19-Feb-1999 #sequence_revision 19-Feb-1999 #text_change 29-Oct-1999

C;Accession: T01702

R;Zheng, D.; Nielsen, B.L.; Daniell, H.

Curr. Genet. 32, 125-131, 1997

A;Title: A 7.5-kbp region of the maize (T cytoplasm) mitochondrial genome contains a ch

A;Reference number: Z14403; MUID:97439744; PMID:9294260

A;Accession: T01702

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-48 <ZHE>

A;Cross-references: EMBL:U42754; NID:g1150864; PIDN:AAB91424.1; PID:g1150866

A;Experimental source: strain cmsF

C;Genetics:

A;Genome: mitochondrion

C;Keywords: mitochondrion

Query Match 58.1%; Score 18; DB 2; Length 48;
Best Local Similarity 80.0%; Pred. No. 2.3e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 RLXER 10
|||
Db 36 RLLEER 40

RESULT 12

S66328

protein kinase AK15 (EC 2.7.1.-) - Arabidopsis thaliana (fragment)

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 24-Sep-1999

C;Accession: S66328; S58263

R;Thuemmler, F.; Kirchner, M.; Teuber, R.; Dittrich, P.

Plant Mol. Biol. 29, 551-565, 1995

A;Title: Differential accumulation of the transcripts of 22 novel protein kinase genes

A;Reference number: S66314; MUID:96123233; PMID:8534852

A;Accession: S66328

A;Molecule type: DNA

A;Residues: 1-58 <THU>

A;Cross-references: EMBL:X86960; NID:g928895; PIDN:CAA60523.1; PID:g928896

C;Genetics:

A;Gene: AK15

A;introns: 47/1

C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homol

C;Keywords: ATP; phosphotransferase; protein kinase
F;1-58/Domains: protein kinase homology (fragment) <KIN>

Query Match 58.1%; Score 18; DB 2; Length 58;
Best Local Similarity 30.0%; Pred. No. 2.8e+02;
Matches 3; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 1 YXXXXRLXER 10
DB 48 YAYTLRIDEX 57

RESULT 13

E84008
Hypothetical protein BH2869 [imported] - Bacillus halodurans (strain C-125)

C;Species: Bacillus halodurans

C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001

C;Accession: E84008

R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hirai

Nucleic Acids Res. 28, 4317-4331, 2000

A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and

A;Reference number: A83650; MUID:20512582; PMID:11058132

A;Accession: E84008

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-59 <STO>

A;Cross-references: GB:AP001516; GB:BA000004; NID:g10175192; PIDN:BA806588.1; GSPDB:GN00

A;Experimental source: strain C-125

C;Genetics:

A;Gene: BH2869

Query Match 58.1%; Score 18; DB 2; Length 59;
Best Local Similarity 80.0%; Pred. No. 2.8e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 6 RLXER 10
DB 47 RLXER 51

RESULT 14

T42259

phosphoprotein phosphatase (EC 3.1.3.16) 1 catalytic chain - Caenorhabditis elegans (firs

C;Species: Caenorhabditis elegans

C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 02-Jun-2000

C;Accession: T42259

R;Zeke, T.; Gergely, P.; Dombradi, V.

submitted to the EMBL Data Library, July 1996

A;Description: The catalytic subunits of Ser/Thr protein phosphatases from Caenorhabditis

A;Reference number: Z22131

A;Accession: T42259

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-24 <ZEK>

A;Cross-references: EMBL:Z77733; PIDN:CAB01292.1

C;Genetics:

A;Note: PPI

C;Superfamily: phosphoprotein phosphatase; phosphoesterase core homology; phosphoprotein

C;Keywords: phosphoric monoester hydrolase

Query Match 54.8%; Score 17; DB 2; Length 24;
Best Local Similarity 44.4%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 YXXXXRLXE 9
DB 2 YSDLRLE 10

RESULT 15

C87422

hypothetical protein CC1394 [imported] - Caulobacter crescentus

C;Species: Caulobacter crescentus

C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001

C;Accession: C87422

R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.

B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon.

n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A;Title: Complete Genome Sequence of Caulobacter crescentus.

A;Reference number: A87249; MUID:21173698; PMID:11259647

A;Accession: C87422

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-39 <STO>

A;Cross-references: GB:AE005673; NID:g13422751; PIDN:AAK23375.1; GSPDB:GN00148

C;Genetics:

A;Gene: CC1394

Query Match 54.8%; Score 17; DB 2; Length 39;
Best Local Similarity 80.0%; Pred. No. 3.7e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 6 RLXER 10
DB 27 RLXER 31

Search completed: December 19, 2002, 16:22:58
Job time : 9.67647 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 19, 2002, 16:02:26 ; Search time 23.8235 Seconds
(without alignments)
55.932 Million cell updates/sec

Title: US-08-653-294C-38
Perfect score: 31
Sequence: 1 REXLRXXXXY 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 473727

Minimum DB seq length: 0
Maximum DB seq length: 60

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	23	74.2	10	16	HLA-B2702 CTL modu
2	23	74.2	10	16	Peptide B2702.75-84(T)
3	23	74.2	10	19	Peptide B2702.75-8
4	23	74.2	14	22	Human Class I HLA-
5	23	74.2	14	22	Human Class I HLA-
6	23	74.2	14	22	Human Class I HLA-
7	23	74.2	14	22	Human Class I HLA-
8	23	74.2	14	22	Human Class I HLA-
9	23	74.2	14	22	Human Class I HLA-
10	23	74.2	20	16	HLA-B2702 CTL modu

11	23	74.2	20	16	AAR92910	HLA-B2702 CTL modu
12	23	74.2	20	19	AAW33792	Peptide B2702.84-7
13	23	74.2	20	19	AAW33793	Peptide B2702.84-7
14	23	74.2	36	22	AAB65129	Gene #26 associate
15	23	74.2	44	21	AAG06983	Arabidopsis thalia
16	22	71.0	10	14	AAR41208	Peptide fragment o
17	22	71.0	10	14	AAR41209	Peptide fragment o
18	22	71.0	10	14	AAR41212	Peptide fragment o
19	22	71.0	10	16	AAR95413	Alphal-helix of HL
20	22	71.0	10	16	AAR95423	HLA-B2705.75-84.
21	22	71.0	10	16	AAR95425	HLA-B2702.75-84(D)
22	22	71.0	10	16	AAR83061	HLA-B7 CTL modul
23	22	71.0	10	16	AAR83062	HLA-B2702 CTL modu
24	22	71.0	10	16	AAR83075	HLA-B2702 CTL modu
25	22	71.0	10	16	AAR83094	HLA-B2702 CTL modu
26	22	71.0	10	17	AAW07512	T-cell modulating
27	22	71.0	10	17	AAW07513	T-cell modulating
28	22	71.0	10	17	AAW07515	T-cell modulating
29	22	71.0	10	19	AAW33784	Peptide B2702.75-8
30	22	71.0	10	19	AAW33785	Peptide B2705.75-8
31	22	71.0	10	19	AAW33786	Peptide B7.75-84 t
32	22	71.0	10	19	AAW33787	Peptide B2702.75-8
33	22	71.0	10	19	AAW33796	Peptide B7.75-84 t
34	22	71.0	10	19	AAW47265	Immunomodulatory p
35	22	71.0	10	19	AAW47269	Immunomodulatory p
36	22	71.0	10	22	AAW72477	Immunosuppressive
37	22	71.0	10	22	AAW72487	Immunosuppressive
38	22	71.0	10	22	AAW72488	Immunosuppressive
39	22	71.0	14	22	AAB59407	Human Class I HLA-
40	22	71.0	14	22	AAB59413	Human Class I HLA-
41	22	71.0	14	22	AAB59414	Human Class I HLA-
42	22	71.0	14	22	AAB59415	Human Class I HLA-
43	22	71.0	14	22	AAB59455	Human Class I HLA-
44	22	71.0	15	16	AAR92912	HLA-B2702 CTL modu
45	22	71.0	15	19	AAW33795	Peptide B2702.70-8

ALIGNMENTS

RESULT 1
AAR95426
ID AAR95426 standard; peptide; 10 AA.
XX
AC AAR95426;
XX
DT 12-NOV-1996 (first entry)
XX
DE HLA-B2702.75-84(T).
XX
HLA; p74; alphal-helix; human-leucocyte-associated antigen; inhibitor;
KW T-cell lyase; membrane protein; mammal; heat shock protein; Hsc70; APC;
KW B cell; calcium influx; cytotoxic T lymphocyte; CTL; differentiation;
KW Cytolysis; antigen presenting cell.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 6 /note= "I6T mutation"
FT
XX
XX WO9513288-A1.
XX
PD 18-MAY-1995.
XX
PF 10-NOV-1994; 94WO-US12985.
XX
PR 10-NOV-1993; 93US-0150493.
XX
PA (STRD) UNIV LELAND STANFORD JUNIOR.
XX
XX Clayberger C, Krensky AM;
XX

DR WPI; 1995-194027/25.

XX Compens. comprising lymphoid surface membrane proteins - which may

PT inhibit cytolytic activity and differentiation of CTLs.

XX Example; Page 11; 29pp; English.

XX AA935413, and AA935415-R95431 represent palindromes and fragments of

CC human-leucocyte-associated antigens. This sequence represents the

CC HLA-B2702.75-84(7). These sequences can be used to isolate the protein

CC p74 from a T-cell lysate. p74 is a T-cell surface membrane protein

CC associated with T-cell activation in mammalian T-cells, and is also

CC immunologically cross reactive with the heat shock protein Hsc70. p74

CC is found in a limited number of cell types, but is particularly expressed

CC on B and T cells. p74 can be isolated by lysis of a suitable cell with

CC an amphoteric detergent, and then passed through an affinity column

CC containing a covalently bound HLA-B2702 palindromic peptide.

CC Compositions comprising the extracellular fragment of p74 combined with

CC HLA-B2702.60-84 (see AA935416), induces calcium influx, and inhibits

CC cytotoxic T lymphocyte (CTL) differentiation or cytotoxicity. Candidate

CC compounds can be screened for their effect on the cytolytic activity of

CC T-cells, by combining them with the extracellular portion of p74 and

CC determining the amount of binding between the candidate compound and

CC p74. Modulation of CTL activity can be inhibited in a cellular

CC composition containing T-cells and antigen presenting cells (APCs), by

CC adding to the mix the extracellular portion of p74, in an amount

CC sufficient to compete with p74 for the binding of the p74 ligand.

CC

XX

SQ Sequence 10 AA;

Query Match 74.2%; Score 23; DB 16; Length 10;

Best Local Similarity 50.0%; Pred. No. 11;

Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 REXLRXXXXY 10

1 REXLRXXXXY 10

DB 1 REXLRXXXXY 10

RESULT 2

AA83095

ID AA83095 standard; peptide; 10 AA.

XX

AC AA83095;

XX

DT 16-MAY-1996 (first entry)

XX

DE HLA-B2702 CTL modulating peptide (B2702.75-84(7)).

XX

KM Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;

KM immunosuppressant; graft versus host disorder; transplantation; therapy;

KM class I MHC; HLA-B2702.

XX

OS Synthetic.

XX

PN WO9526979-A1.

XX

PD 12-OCT-1995.

XX

PF 05-APR-1995; 95WO-US04349.

XX

PR 05-APR-1994; 94US-0222851.

XX

PA (STRD) UNIV LELAND STANFORD JUNIOR.

XX

PI Clayberger C, Krensky AM, Parham P;

XX

DR WPI; 1995-358582/46.

XX

PT Extension of acceptance period of transplants from MHC unmatched

PT donor hosts - using Class I B75-84 MHC antigen of the recipient

PT host

XX

PS Example 14; Page 34; 80pp; English.

XX

CC AA83061-R83085, AA83090-R83096 and AA832907-R92913 represent fragments

CC of class I major histocompatibility complex (MHC) antigens. This

CC sequence corresponds to residues 75-84 of the alpha-1 domain of the class

CC I MHC HLA-B2702. This sequence showed no inhibitory effect upon

CC cytotoxic T lymphocytes (CTLs). These sequences can be used to extend

CC the period of acceptance by a recipient of a transplant from an MHC

CC unmatched donor. The peptides are administered to a patient in

CC conjunction with a subtherapeutic amount of an immunosuppressant. This

CC is administered to the patient for a limited period of time (compared to

CC the lifetime administration for current treatments). The peptides

CC particularly modulate (or inhibit) the activity of the CTLs of the

CC patient.

XX

SQ Sequence 10 AA;

Query Match 74.2%; Score 23; DB 16; Length 10;

Best Local Similarity 50.0%; Pred. No. 11;

Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 REXLRXXXXY 10

1 REXLRXXXXY 10

DB 1 REXLRXXXXY 10

RESULT 3

AAW33788

ID AAW33788 standard; peptide; 10 AA.

XX

AC AAW33788;

XX

DT 19-JUN-1998 (first entry)

XX

DE Peptide B2702.75-84(780) tested for immunomodulating activity.

XX

KM Immunomodulating dimer; immunosuppressant drug; CTL activation;

KM transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;

KM rejection.

XX

OS Synthetic.

XX

PN WO9744351-A1.

XX

PD 27-NOV-1997.

XX

PF 22-MAY-1997; 97WO-US08689.

XX

PR 24-MAY-1996; 96US-0653294.

XX

PA (STRD) UNIV LELAND STANFORD JUNIOR.

XX

PI Beulow R, Clayberger C, Krensky AM;

XX

DR WPI; 1998-086530/08.

XX

PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B

PT alpha-1 domain, used for preventing rejection of transplants or

PT treating autoimmune diseases

XX

PS Example 1; Page 19; 41pp; English.

XX

CC Peptides AAW33784-98 and AAW33778-9 were assayed for their

CC immunomodulating activity. A peptide-type compound or variant is claimed

CC which has immunomodulating activity, including the N-terminal acylated

CC and/or C-terminal amidated or esterified forms of up to 60 amino acids,

CC where the peptide-type compound comprises the formula: A-B, where A, B =

CC (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or V; aa77 =

CC D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a hydrophobic or

CC small amino acid; aa82 = R or L; aa83 = G or R; and aa represents amino

CC acid. The sequence in the brackets may optionally be absent or truncated

CC at any peptide type bond within the brackets. The compounds comprise

CC amino acid sequences related to a Class I HLA-B alpha1 domain (positions
 CC 79-84). They can be used to inhibit cytotoxic T-lymphocytes (CTL) from
 CC undesirably attacking cells in a host or in vitro. They can also be
 CC used in combination with antigenic peptides or proteins of interest to
 CC activate CTLs. They can also inhibit the proliferation of T cells in
 CC response to anti-CD3. The peptide can be used for preventing rejection
 CC of transplants or for treating autoimmune diseases, e.g. diabetes,
 CC rheumatoid arthritis and lupus erythematosus. The products can also be
 CC used for detection and diagnosis.

XX Sequence 10 AA;

Query Match 74.2%; Score 23; DB 19; Length 10;
 Best Local Similarity 50.0%; Pred. No. 11;
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 REXLRXXXXY 10
 |||||
 Db 1 RENLRALRY 10

RESULT 4

AAB59405
 ID AAB59405 standard; Peptide; 14 AA.

XX AAB59405;

DT 22-MAR-2001 (first entry)

DE Human Class I HLA-B alpha1-domain-derived peptide #5.

KW Human; human leukocyte antigen-B; HLA-B; immunosuppressive;
 KW cytotoxic T cell activation inhibition; cell transplantation;
 KW natural killer cell activation inhibition; organ transplantation.

XX Homo sapiens.

OS Synthetic.

XX US6162434-A.

PD 19-DEC-2000.

PF 03-MAY-1995; 95US-0433613.

XX 03-MAY-1995; 95US-0433613.

PR (SANG-) SANGSTAT MEDICAL CORP.

PA Buelow R;

PI WPI; 2001-111720/12.

DR Inhibiting activation of cytotoxic T lymphocytes and/or natural killer
 PT cells involves combining cells with compound comprising oligopeptide
 PT with contiguous sequence of human leukocyte antigen-B alpha1 domain -

PS Disclosure; Column 7-8; 20pp; English.

XX The present sequence is a peptide which may be used in a method for
 CC inhibiting activation of cytotoxic T lymphocytes and/or natural killer
 CC cells. The method involves combining the cells with a compound
 CC comprising an oligopeptide of at least 6 amino acids comprising a
 CC contiguous sequence of a human leukocyte antigen (HLA)-B alpha1 domain
 CC including amino acids 84-86, where amino acids 84-86 are YYW.
 CC The method is useful for transplanting a donor mammalian organ or cells
 CC to a mammalian recipient. The organ or cells are combined with the
 CC compound prior to implanting in the mammalian recipient, or the compound
 CC may be administered to the mammalian recipient in a period extending
 CC from prior to implanting the donor organ or cells.

XX Sequence 14 AA;

Query Match 74.2%; Score 23; DB 22; Length 14;

Best Local Similarity 50.0%; Pred. No. 16;
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 REXLRXXXXY 10
 |||||
 Db 5 RENLRALRY 14

RESULT 5

AAB59410
 ID AAB59410 standard; Peptide; 14 AA.

XX AAB59410;

DT 22-MAR-2001 (first entry)

DE Human Class I HLA-B alpha1-domain-derived peptide #10.

KW Human; human leukocyte antigen-B; HLA-B; immunosuppressive;
 KW cytotoxic T cell activation inhibition; cell transplantation;
 KW natural killer cell activation inhibition; organ transplantation.

XX Homo sapiens.

OS Synthetic.

XX US6162434-A.

PD 19-DEC-2000.

PF 03-MAY-1995; 95US-0433613.

XX 03-MAY-1995; 95US-0433613.

PR (SANG-) SANGSTAT MEDICAL CORP.

PA Buelow R;

PI WPI; 2001-111720/12.

DR Inhibiting activation of cytotoxic T lymphocytes and/or natural killer
 PT cells involves combining cells with compound comprising oligopeptide
 PT with contiguous sequence of human leukocyte antigen-B alpha1 domain -

PS Disclosure; Column 7-8; 20pp; English.

XX The present sequence is a peptide which may be used in a method for
 CC inhibiting activation of cytotoxic T lymphocytes and/or natural killer
 CC cells. The method involves combining the cells with a compound
 CC comprising an oligopeptide of at least 6 amino acids comprising a
 CC contiguous sequence of a human leukocyte antigen (HLA)-B alpha1 domain
 CC including amino acids 84-86, where amino acids 84-86 are YYW.
 CC The method is useful for transplanting a donor mammalian organ or cells
 CC to a mammalian recipient. The organ or cells are combined with the
 CC compound prior to implanting in the mammalian recipient, or the compound
 CC may be administered to the mammalian recipient in a period extending
 CC from prior to implanting the donor organ or cells.

XX Sequence 14 AA;

Query Match 74.2%; Score 23; DB 22; Length 14;
 Best Local Similarity 50.0%; Pred. No. 16;
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 REXLRXXXXY 10
 |||||
 Db 5 RSLRLALRY 14

RESULT 6

AAB59436
 ID AAB59436 standard; Peptide; 14 AA.

XX AAB59436;

XX	22-MAR-2001	(first entry)	
DT			
XX			
DE	Human Class I HLA-B alpha1-domain-derived peptide #36.		
XX			
XX			
KM	Human; human leukocyte antigen-B; HLA-B; immunosuppressive;		
KW	cytotoxic T cell activation inhibition; cell transplantation;		
KM	natural killer cell activation inhibition; organ transplantation.		
XX			
OS	Homo sapiens.		
OS	Synthetic.		
XX			
PN	US6162434-A.		
XX			
PD	19-DEC-2000.		
XX			
PF	03-MAY-1995; 95US-0433613.		
XX			
PR	03-MAY-1995; 95US-0433613.		
XX			
PA	(SANG-) SANGSTAT MEDICAL CORP.		
XX			
PI	Buelow R;		
XX			
DR	WPI; 2001-111720/12.		
XX			
PT	Inhibiting activation of cytotoxic T lymphocytes and/or natural killer		
PT	cells involves combining cells with compound comprising oligopeptide		
PT	with contiguous sequence of human leukocyte antigen-B alpha1 domain		
XX			
PS	Disclosure; Column 9-10; 20pp; English.		
XX			
CC	The present sequence is a peptide which may be used in a method for		
CC	inhibiting activation of cytotoxic T lymphocytes and/or natural killer		
CC	cells. The method involves combining the cells with a compound		
CC	comprising an oligopeptide of at least 6 amino acids comprising a		
CC	contiguous sequence of a human leukocyte antigen (HLA)-B alpha1 domain		
CC	including amino acids 84-86, where amino acids 84-86 are YW.		
CC	The method is useful for transplanting a donor mammalian organ or cells		
CC	to a mammalian recipient. The organ or cells are combined with the		
CC	compound prior to implanting in the mammalian recipient, or the compound		
CC	may be administered to the mammalian recipient in a period extending		
CC	from prior to implanting the donor organ or cells.		
XX			
SQ	Sequence 14 AA;		
	Query Match 74.2%; Score 23; DB 22; Length 14;		
	Best Local Similarity 50.0%; Pred. No. 16;		
	Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;		
OY	1 REXLRXXXXY 10		
	5 REALRLRLRY 14		
DB			
	RESULT 7		
	AAB59439		
ID	AAB59439 standard; peptide; 14 AA.		
XX			
AC	AAB59439;		
XX			
DT	22-MAR-2001 (first entry)		
XX			
DE	Human Class I HLA-B alpha1-domain-derived peptide #39.		
XX			
KM	Human; human leukocyte antigen-B; HLA-B; immunosuppressive;		
KW	cytotoxic T cell activation inhibition; cell transplantation;		
KM	natural killer cell activation inhibition; organ transplantation.		
XX			
OS	Homo sapiens.		
OS	Synthetic.		
XX			
PN	US6162434-A.		

[illegible]

PT Inhibiting activation of cytotoxic T lymphocytes and/or natural killer
 PT cells involves combining cells with compound comprising oligopeptide
 PT with contiguous sequence of human leukocyte antigen-B alaphal domain -
 XX Disclosure; Column 9-10; 20pp; English.

XX The present sequence is a peptide which may be used in a method for
 CC inhibiting activation of cytotoxic T lymphocytes and/or natural killer
 CC cells. The method involves combining the cells with a compound
 CC comprising an oligopeptide of at least 6 amino acids comprising a
 CC contiguous sequence of a human leukocyte antigen (HLA)-B alaphal domain
 CC including amino acids 84-86, where amino acids 84-86 are YVW.
 CC The method is useful for transplanting a donor mammalian organ or cells
 CC to a mammalian recipient. The organ or cells are combined with the
 CC compound prior to implanting in the mammalian recipient, or the compound
 CC may be administered to the mammalian recipient in a period extending
 CC from prior to implanting the donor organ or cells.

XX Sequence 14 AA;

Query Match 74.2%; Score 23; DB 22; Length 14;
 Best Local Similarity 50.0%; Pred. No. 16;
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 REXLRXXXXY 10
 |||||
 Db 5 REXLRIAARY 14

RESULT 9

AAB59441
 ID AAB59441 standard; Peptide; 14 AA.

AC AAB59441;

XX 22-MAR-2001 (first entry)

DE Human Class I HLA-B alaphal-domain-derived peptide #41.

XX Human; human leukocyte antigen-B; HLA-B; immunosuppressive;

KW cytotoxic T cell activation inhibition; cell transplantation;
 KW natural killer cell activation inhibition; organ transplantation.

XX Homo sapiens.

OS Synthetic.

XX US6162434-A.

XX 19-DEC-2000.

XX 03-MAY-1995; 95US-0433613.

XX 03-MAY-1995; 95US-0433613.

XX (SANG-) SANGSTAT MEDICAL CORP.

XX Buelow R;

XX WPI; 2001-111720/12.

XX Inhibiting activation of cytotoxic T lymphocytes and/or natural killer
 PT cells involves combining cells with compound comprising oligopeptide
 PT with contiguous sequence of human leukocyte antigen-B alaphal domain -

XX Disclosure; Column 9-10; 20pp; English.

XX The present sequence is a peptide which may be used in a method for
 CC inhibiting activation of cytotoxic T lymphocytes and/or natural killer
 CC cells. The method involves combining the cells with a compound
 CC comprising an oligopeptide of at least 6 amino acids comprising a
 CC contiguous sequence of a human leukocyte antigen (HLA)-B alaphal domain
 CC including amino acids 84-86, where amino acids 84-86 are YVW.
 CC The method is useful for transplanting a donor mammalian organ or cells

CC to a mammalian recipient. The organ or cells are combined with the
 CC compound prior to implanting in the mammalian recipient, or the compound
 CC may be administered to the mammalian recipient in a period extending
 CC from prior to implanting the donor organ or cells.

SQ Sequence 14 AA;

Query Match 74.2%; Score 23; DB 22; Length 14;
 Best Local Similarity 50.0%; Pred. No. 16;
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 REXLRXXXXY 10
 |||||
 Db 5 REXLRIALAY 14

RESULT 10

AAR92909

ID AAR92909 standard; peptide; 20 AA.

XX AAR92909;

XX 16-MAY-1996 (first entry)

DE HLA-B2702 CTL modulating peptide (B2702.84-75/75-84 (T)).

XX Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;
 KW immunosuppressant; Graft versus host disorder; transplantation; therapy;
 KW class I MHC; HLA-B2702.

XX Synthetic.

XX WO9526979-A1.

XX 12-OCT-1995.

XX 05-APR-1995; 95WO-US04349.

XX 05-APR-1994; 94US-0222851.

XX (STRD) UNIV LELAND STANFORD JUNIOR.

XX Clayberger C, Krensky AM, Parham P;

XX WPI; 1995-358582/46.

XX Extension of acceptance period of transplants from MHC unmatched
 PT donor hosts - using Class I B75-84 MHC antigen of the recipient
 PT host

XX Example 15; Page 36; 80pp; English.

XX AAR93061-R83085, AAR93090-R83096 and AAR92907-R92913 represent fragments
 CC of class I major histocompatibility complex (MHC) antigens. This
 CC sequence is an inverted dimer of residues 75-84 of the alpha-1 domain of
 CC the class I MHC HLA-B2702. These sequences can be used to extend the
 CC period of acceptance by a recipient of a transplant from an MHC unmatched
 CC donor. The peptides are administered to a patient in conjunction with a
 CC subtherapeutic amount of an immunosuppressant. This is administered to
 CC the patient for a limited period of time (compared to the lifetime
 CC administration for current treatments). The peptides particularly
 CC modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)
 CC of the patient.

SQ Sequence 20 AA;

Query Match 74.2%; Score 23; DB 16; Length 20;
 Best Local Similarity 50.0%; Pred. No. 22;
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 REXLRXXXXY 10
 |||||
 Db 11 REXLRALRY 20

RESULT 11
 AAR92910 standard; peptide; 20 AA.
 ID AAR92910;
 AC AAR92910;
 XX 16-MAY-1996 (first entry)
 DT HLA-B2702 CTL modulating peptide (B2702.84-75(T)/75-84(T)).
 DE
 XX Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;
 KW immunosuppressant; graft versus host disorder; transplantation; therapy;
 KW class I MHC; HLA-B2702.
 XX Synthetic.
 OS WO9526979-A1.
 PN 12-OCT-1995.
 PD 05-APR-1995; 95WO-US04349.
 PF 05-APR-1994; 94US-0222851.
 PR (STRD) UNIV LELAND STANFORD JUNIOR.
 XX Clayberger C, Krensky AM, Parham P;
 PI WPI; 1995-358582/46.
 DR Extension of acceptance period of transplants from MHC unmatched
 XX donor hosts - using Class I B75-84 MHC antigen of the recipient
 PT host
 PS Example 15; Page 36; 80pp; English.
 XX AAR83061-R83085, AAR83090-R83096 and AAR92907-R92913 represent fragments
 CC of class I major histocompatibility complex (MHC) antigens. This
 CC sequence is an inverted dimer of residues 75-84 of the alpha-1 domain of
 CC the class I MHC HLA-B2702. These sequences can be used to extend the
 CC period of acceptance by a recipient of a transplant from an MHC unmatched
 CC donor. The peptides are administered to a patient in conjunction with a
 CC subtherapeutic amount of an immunosuppressant. This is administered to
 CC the patient for a limited period of time (compared to the lifetime
 CC administration for current treatments). The peptides particularly
 CC modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)
 CC of the patient.
 CC
 XX Sequence 20 AA;
 SQ
 Query Match 74.2%; Score 23; DB 16; Length 20;
 Best Local Similarity 50.0%; Pred. No. 22;
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 OY 1 REXLRXXXXY 10
 |||||
 DB 11 RENLRTALRY 20
 |||||
 RESULT 12
 AAW33792 standard; peptide; 20 AA.
 ID AAW33792;
 AC AAW33792;
 XX 19-JUN-1998 (first entry)
 DT Peptide B2702.84-75/75-84T tested for immunomodulating activity.
 DE
 XX Immunomodulating dimer; immunosuppressant drug; CTL activation;
 KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;
 KW

KW rejection.
 XX Synthetic.
 OS Homo sapiens.
 XX WO9744351-A1.
 PN 27-NOV-1997.
 PD 22-MAY-1997; 97WO-US08689.
 PF 24-MAY-1996; 96US-0653294.
 PR (STRD) UNIV LELAND STANFORD JUNIOR.
 XX Beulow R, Clayberger C, Krensky AM;
 PI WPI; 1998-086530/08.
 DR New immunomodulating dimer peptide(s) - based on a Class I HLA-B
 PT alpha-1 domain, used for preventing rejection of transplants or
 PT treating autoimmune diseases
 PS Example 1; Page 19; 41pp; English.
 XX Peptides AAW33784-98 and AAW33778-9 were assayed for their
 CC immunomodulating activity. A peptide-type compound or variant is claimed
 CC which has immunomodulating activity, including the N-terminal acylated
 CC and/or C-terminal amidated or esterified forms of up to 60 amino acids,
 CC where the peptide-type compound comprises the formula: A-B, where A, B =
 CC (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or V; aa77 =
 CC D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a hydrophobic or
 CC small amino acid; aa82 = R or L; aa83 = G or R; and aa represents amino
 CC acid. The sequence in the brackets may optionally be absent or truncated
 CC at any peptide type bond within the brackets. The compounds comprise
 CC amino acid sequences related to a Class I HLA-B alpha domain (positions
 CC 79-84). They can be used to inhibit cytotoxic T-lymphocytes (CTL) from
 CC undesirably attacking cells in a host or in vitro. They can also be
 CC used in combination with antigenic peptides or proteins of interest to
 CC activate CTLs. They can also inhibit the proliferation of T cells in
 CC response to anti-CD3. The peptide can be used for preventing rejection
 CC of transplants or for treating autoimmune diseases, e.g. diabetes,
 CC rheumatoid arthritis and lupus erythematosus. The products can also be
 CC used for detection and diagnosis.
 CC
 XX Sequence 20 AA;
 SQ
 Query Match 74.2%; Score 23; DB 19; Length 20;
 Best Local Similarity 50.0%; Pred. No. 22;
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 OY 1 REXLRXXXXY 10
 |||||
 DB 11 RENLRTALRY 20
 |||||
 RESULT 13
 AAW33793 standard; peptide; 20 AA.
 ID AAW33793;
 AC AAW33793;
 XX 19-JUN-1998 (first entry)
 DT Peptide B2702.84-75T/75-84T tested for immunomodulating activity.
 DE
 XX Immunomodulating dimer; immunosuppressant drug; CTL activation;
 KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;
 KW rejection.
 KW Synthetic.
 OS Homo sapiens.
 XX

PN W09744351-A1.
 XX 27-NOV-1997.
 XX 22-MAY-1997; 97WO-US08689.
 XX 24-MAY-1996; 96US-0653294.
 XX (STRD) UNIV LELAND STANFORD JUNIOR.
 PA Beulow R, Clayberger C, Krensky AM;
 XX WPI; 1998-086530/08.
 XX New immunomodulating dimer peptide(s) - based on a Class I HLA-B
 PT alpha-1 domain, used for preventing rejection of transplants or
 PT treating autoimmune diseases
 XX Example 1; Page 19; 41pp; English.
 XX Peptides AAW33784-98 and AAW33778-9 were assayed for their
 CC immunomodulating activity. A peptide-type compound or variant is claimed
 CC which has immunomodulating activity, including the N-terminal acylated
 CC and/or C-terminal amidated or esterified forms of up to 60 amino acids,
 CC where the peptide-type compound comprises the formula; A-B, where A, B =
 CC (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or V; aa77 =
 CC D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a hydrophobic or
 CC small amino acid; aa82 = R or L; aa83 = G or R; and aa represents amino
 CC acid. The sequence in the brackets may optionally be absent or truncated
 CC at any peptide type bond within the brackets. The compounds comprise
 CC amino acid sequences related to a Class I HLA-B alpha1 domain (positions
 CC 79-84). They can be used to inhibit cytotoxic T-lymphocytes (CTL) from
 CC undesirably attacking cells in a host or in vitro. They can also be
 CC used in combination with antigenic peptides or proteins of interest to
 CC activate CTLs. They can also inhibit the proliferation of T cells in
 CC response to anti-CD3. The peptide can be used for preventing rejection
 CC of transplants or for treating autoimmune diseases, e.g. diabetes,
 CC rheumatoid arthritis and lupus erythematosus. The products can also be
 CC used for detection and diagnosis.
 XX Sequence 20 AA;
 SQ
 Query Match 74.2%; Score 23; DB 19; Length 20;
 Best Local Similarity 50.0%; Pred. No. 22;
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 1 REXLRXXXXY 10
 Db 11 REXLRXXXXY 20
 RESULT 14
 AAB65129
 ID AAB65129 standard; peptide; 36 AA.
 AC AAB65129;
 XX 23-MAR-2001 (first entry)
 DT Gene #26 associated peptide #20.
 DE Secreted protein; gene therapy; vaccine; cancer; leukemia;
 KW autoimmune disease; allergy; inflammation; graft rejection;
 KW hyperproliferation; cardiovascular; infection.
 XX Homo sapiens.
 OS
 XX WO200075375-A1.
 PN 14-DEC-2000.
 XX 02-JUN-2000; 2000WO-US15187.
 PF
 XX

PR 07-JUN-1999; 99US-0137725.
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA Ruben SM, Birse CE, Duan RD, Soppet DR, Rosen CA, Shi Y;
 PI Lafleur DW, Olsen HS, Ebner R, Florence KA, Ni J, Young PE;
 XX WPI; 2001-061741/07.
 DR Nucleic acids encoding 26 human secreted polypeptides, useful e.g. for
 XX preventing, diagnosing and/or treating cancers and for promoting wound
 PT healing -
 PT Disclosure; Page 81; 530pp; English.
 PS The present invention relates to 26 secreted human proteins. The
 XX proteins may be used in the prevention, diagnosis and treatment of
 CC diseases associated with inappropriate polypeptide expression.
 CC For example, they may be used in gene therapy or in vaccines.
 CC Typical of diseases which are potentially treatable are cancers
 CC (including leukemia), autoimmune diseases, allergies, inflammation,
 CC graft rejection, hyperproliferation, cardiovascular diseases
 CC (particularly critical limb ischemia and coronary disease) and any
 CC involving abnormal angiogenesis, neurodegeneration and/or
 CC infectious diseases.
 XX Sequence 36 AA;
 SQ
 Query Match 74.2%; Score 23; DB 22; Length 36;
 Best Local Similarity 50.0%; Pred. No. 40;
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 1 REXLRXXXXY 10
 Db 11 REXLRXXXXY 20
 RESULT 15
 AAG06983
 ID AAG06983 standard; Protein; 44 AA.
 XX AAG06983;
 AC AAG06983;
 XX 17-OCT-2000 (first entry)
 DT Arabidopsis thaliana protein fragment SEQ ID NO: 3960.
 DE Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 XX Arabidopsis thaliana.
 OS
 XX EP1033405-A2.
 PN 06-SEP-2000.
 PD 25-FEB-2000; 2000EP-0301439.
 PF 25-FEB-1999; 99US-0121825.
 XX 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 23-APR-1999; 99US-0130891.

PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0135621.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.

PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151086.
PR 27-AUG-1999; 99US-0151086.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152353.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154479.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.

PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 74.2%; Score 23; DB 21; Length 44;
Best Local Similarity 50.0%; Pred. No. 48;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 REXLXXXXY 10
Db 20 RETLRNLNY 29

Search completed: December 19, 2002, 16:19:05
Job time : 23.8235 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 19, 2002, 16:14:32 ; Search time 17.6471 Seconds
(without alignments)
116.760 Million cell updates/sec

Title: US-08-653-294C-38

Perfect score: 31

Sequence: 1 REXLRXXXYY 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 45785

Minimum DB seq length: 0

Maximum DB seq length: 60

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_21.*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	23	74.2	52	11 Q9CULO	Q9cul0 mus musculus
2	22	71.0	39	7 O19688	O19688 homo sapien
3	22	71.0	48	7 Q9TQC5	Q9tqc5 homo sapien
4	22	71.0	48	7 Q9TQC2	Q9tqc2 homo sapien
5	21	67.7	51	5 Q24623	Q24623 drosophila
6	19	61.3	46	7 Q8WM99	Q8wm99 homo sapien
7	19	61.3	47	7 O19787	O19787 homo sapien
8	19	61.3	57	16 Q8YNC6	Q8ync6 anabaena sp
9	19	61.3	60	10 Q8VYX6	Q8vyx6 pisum sativ
10	18	58.1	18	11 Q9QZX5	Q9qzx5 mus musculus
11	18	58.1	18	11 Q9QZX4	Q9qzx4 mus spretus
12	18	58.1	21	5 Q95SN4	Q95sn4 drosophila
13	18	58.1	26	12 Q83274	Q83274 cucumber mo
14	18	58.1	27	12 Q9J5W8	Q9j5w8 hepatitis c
15	18	58.1	30	10 Q9S871	Q9s871 monordica c
16	18	58.1	31	12 O73430	O73430 human papil

17	18	58.1	33	12 Q90085	Q90085 human papil
18	18	58.1	33	16 Q8USM4	Q8usm4 agrobacteri
19	18	58.1	34	6 Q95LA6	Q95la6 sus scrofa
20	18	58.1	34	15 P90368	P90368 human immun
21	18	58.1	37	16 Q8Z4C8	Q8z4c8 salmonella
22	18	58.1	39	15 O74264	O74264 human immun
23	18	58.1	41	2 Q9AFK5	Q9afk5 shigella fl
24	18	58.1	42	2 Q52153	Q52153 escherichia
25	18	58.1	42	5 Q9NMP5	Q9nmp5 leishmania
26	18	58.1	48	2 Q48715	Q48715 lactococcus
27	18	58.1	50	16 Q9JVI7	Q9jvi7 neisseria m
28	18	58.1	50	17 Q9HS53	Q9hs53 halobacteri
29	18	58.1	51	4 Q9BXP7	Q9bxf7 homo sapien
30	18	58.1	54	2 Q9R1K0	Q9r1k0 streptococc
31	18	58.1	54	12 Q89918	Q89918 cucumber mo
32	18	58.1	54	17 Q96XL7	Q96xl7 sulfolobus
33	18	58.1	56	16 Q8U502	Q8u5q2 agrobacteri
34	18	58.1	57	5 Q9U0S8	Q9u0s8 moniezia ex
35	18	58.1	57	16 Q9K7H8	Q9k7h8 bacillus ha
36	18	58.1	59	16 Q8XLN8	Q8xln8 clostridium
37	17	54.8	20	4 Q9NZ28	Q9nz28 homo sapien
38	17	54.8	20	4 Q9NZ27	Q9nz27 homo sapien
39	17	54.8	25	11 Q9QVL8	Q9qvl8 mus sp. glu
40	17	54.8	25	16 Q8X3V1	Q8x3v1 escherichia
41	17	54.8	27	12 Q9IKF9	Q9ikf9 hepatitis c
42	17	54.8	35	11 Q63896	Q63896 mus sp. nk-
43	17	54.8	35	15 Q71527	Q71527 human immun
44	17	54.8	35	15 Q40494	Q40494 human immun
45	17	54.8	35	15 Q97869	Q97869 human immun

ALIGNMENTS

RESULT 1

ID	Q9CULO	PRELIMINARY;	PRT;	52 AA.
AC	Q9CULO;			
DT	01-JUN-2001 (Tremblrel. 17, Created)			
DT	01-JUN-2001 (Tremblrel. 17, Last sequence update)			
DT	01-MAR-2002 (Tremblrel. 20, Last annotation update)			
DE	653040A22Rik protein (Fragment).			
GN	653040A22Rik.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=C57BL/6J; TISSUE=TESTIS;			
RX	MEDLINE=21085660; PubMed=11217851;			
RA	Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,			
RA	Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,			
RA	Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,			
RA	Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,			
RA	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,			
RA	Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,			
RA	Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,			
RA	Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,			
RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,			
RA	Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,			
RA	Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,			
RA	Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,			
RA	Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,			
RA	Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,			
RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,			
RA	Suzuki H., Toyo-Oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,			
RA	Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,			
RA	Hayashizaki Y.;			
RT	"Functional annotation of a full-length mouse cDNA collection."			
RL	Nature 409:685-690(2001).			
DR	EMBL; AK015634; BAB29909.1; -			
DR	MGI; MGI:1915046; 653040A22Rik.			

DR InterPro; IPRO01230; Prenyl site.
 DR PROSITE; PS00294; PRENYLATION; UNKNOWN_1.
 FT NON TER 1
 SQ SEQUENCE 52 AA; 5506 MW; 0275B42B0A9F900 CRC64;

Query Match 74.2%; Score 23; DB 11; Length 52;
 Best Local Similarity 50.0%; Pred. No. 38;
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 REXLXXXXXX 10
 DB 23 REGURAKRAY 32

RESULT 2

O19688 PRELIMINARY; PRT; 39 AA.

ID O19688
 AC O19688;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE HLA-B*27 variant exon 2 (Alpha domain) (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;

RP SEQUENCE FROM N.A.
 RA Blaszczak R., Weber M., Salama A.;
 RL Submitted (JUN-1995) to the EMBL/Genbank/DBJ databases.
 DR EMBL; X83727; CAA58698.1; -
 DR InterPro; IPR001039; MHC_I.
 DR Pfam; PF00129; MHC_I; 1.
 DR ProDom; PD000050; MHC_I; 1.
 FT NON TER 1
 FT NON TER 39
 SQ SEQUENCE 39 AA; 4748 MW; F5E3FD2A53138908 CRC64;

Query Match 71.0%; Score 22; DB 7; Length 39;
 Best Local Similarity 50.0%; Pred. No. 55;
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 REXLXXXXXX 10
 DB 24 REDLRTLLRY 33

RESULT 3

O970C5 PRELIMINARY; PRT; 48 AA.

ID O970C5
 AC O970C5;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE MHC class I HLA-B (Fragment).
 GN HLA-B.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;

RP SEQUENCE FROM N.A.
 RA Adams B., McLean I.L., Roberts M., Woodfield D.G.;
 RL Submitted (JUN-1997) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AF011546; AAD46382.2; -
 DR InterPro; IPR001039; MHC_I.
 DR Pfam; PF00129; MHC_I; 1.
 DR ProDom; PD000050; MHC_I; 1.
 FT NON TER 1
 FT NON TER 48
 SQ SEQUENCE 48 AA; 5841 MW; DFBEF6652D5B60C3 CRC64;

Query Match 71.0%; Score 22; DB 7; Length 48;

Best Local Similarity 50.0%; Pred. No. 67;
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 REXLXXXXXX 10
 DB 33 RESLNLNGRY 42

RESULT 4

O970C2 PRELIMINARY; PRT; 48 AA.

ID O970C2
 AC O970C2;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE MHC class I HLA-B (Fragment).
 GN HLA-B.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;

RP SEQUENCE FROM N.A.
 RA Adams B., McLean I.L., Roberts M., Woodfield D.G.;
 RL Submitted (JUN-1997) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AF011546; AAD46385.2; -
 DR InterPro; IPR001039; MHC_I.
 DR Pfam; PF00129; MHC_I; 1.
 DR ProDom; PD000050; MHC_I; 1.
 FT NON TER 1
 FT NON TER 48
 SQ SEQUENCE 48 AA; 5841 MW; DFBEF6652D5B60C3 CRC64;

Query Match 71.0%; Score 22; DB 7; Length 48;
 Best Local Similarity 50.0%; Pred. No. 67;
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 REXLXXXXXX 10
 DB 33 RESLNLNGRY 42

RESULT 5

O24623 PRELIMINARY; PRT; 51 AA.

ID O24623
 AC O24623;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE LETHAL(3)87DF protein (Fragment).
 GN L(3)87DF.
 OS Drosophila pseudoobscura (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 CC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 CC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7237;

RP SEQUENCE FROM N.A.
 RA MEDLINE=89158785; PubMed=2493563;
 RA Riley M.A.;
 RT "Nucleotide sequence of the xdh region in Drosophila pseudoobscura and
 an analysis of the evolution of synonymous codons."
 RL Mol. Biol. Evol. 6:333-52(1989).
 DR EMBL; M33977; ABA29021.1; -
 DR FlyBase; FBgn0012726; Dpse1(3)87DF.
 FT NON TER 1
 FT NON TER 51
 SQ SEQUENCE 51 AA; 6212 MW; CS023140B744A005 CRC64;

Query Match 67.7%; Score 21; DB 5; Length 51;
 Best Local Similarity 40.0%; Pred. No. 1.4e+02;
 Matches 4; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 1 REXLXXXXXX 10

Db 28 REAMRQALY 37
|| :|
|| :|

RESULT 6

Q8WM99 PRELIMINARY; PRT; 46 AA.
AC Q8WM99;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE MHC class I (Fragment).
GN HLA-B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BLOOD;
RA Arnaiz-Villena A., Martinez-Laso J., Gomez-Casado E., Bedoya C.,
RA Montoya F.;
RT "HLA-B new alleles in South American Indians.";
RT Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; L36980; AAL31360.1; -;
DR InterPro; IPR001039; MHC.I.
DR Pfam; PF00129; MHC I; 1.
DR ProDom; PD000050; MHC I; 1.
FT NON_TER 1
FT NON_TER 46 46
SQ SEQUENCE 46 AA; 5500 MW; 62ACB787B66D8F4 CRC64;

Query Match 61.3%; Score 19; DB 7; Length 46;
Best Local Similarity 80.0%; Pred. No. 4.5e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 REXLR 5
|| :|
|| :|
Db 40 RESLR 44

RESULT 7

O19787 PRELIMINARY; PRT; 47 AA.
AC O19787;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE HLA class I antigen (Fragment).
GN HLA-B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94090601; PubMed=7903491;
RA Yoshida M., Kimura A., Katsuragi K., Numano F., Sasazuki T.;
RT "DNA typing of HLA-B gene in Takayasu's arteritis.";
RL Tissue Antigens 42:87-90(1993).
DR EMBL; S67638; AAB29305.2; -;
DR InterPro; IPR001039; MHC.I.
DR Pfam; PF00129; MHC I; 1.
DR ProDom; PD000050; MHC I; 1.
FT NON_TER 1
FT NON_TER 47 47
SQ SEQUENCE 47 AA; 5608 MW; 30AF5787B6669B02 CRC64;

Query Match 61.3%; Score 19; DB 7; Length 47;
Best Local Similarity 80.0%; Pred. No. 4.6e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 REXLR 5

Db 41 RESLR 45
|| :|
|| :|

RESULT 8

Q8YNC6 PRELIMINARY; PRT; 57 AA.
AC Q8YNC6;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein Asl4640.
GN Asl4640.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimpō S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
cyanobacterium Anabaena sp. strain PCC 7120.";
RL DNA Res. 8:205-213(2001).
DR EMBL; AP003597; BAB76339.1; -;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 57 AA; 6401 MW; 2F552487C9BC7CD6 CRC64;

Query Match 61.3%; Score 19; DB 16; Length 57;
Best Local Similarity 80.0%; Pred. No. 5.5e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 REXLR 5
|| :|
|| :|
Db 22 RESLR 26

RESULT 9

Q8VYX6 PRELIMINARY; PRT; 60 AA.
AC Q8VYX6;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Ultraviolet-B-inducible auxin-related protein (Fragment).
OS Pisum sativum (Garden pea).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Pisum.
OX NCBI_TaxID=3888;
RN [1]
RP SEQUENCE FROM N.A.
RA Liu L., White M.J., MacRae T.H.;
RT "Identification of ultraviolet-B responsive genes in the pea, Pisum
sativum.";
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY065659; AAL50319.1; -;
FT NON_TER 1
FT NON_TER 60 60
SQ SEQUENCE 60 AA; 6303 MW; 5EB1AA1CF3868B98 CRC64;

Query Match 61.3%; Score 19; DB 10; Length 60;
Best Local Similarity 80.0%; Pred. No. 5.9e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 REXLR 5
|| :|
|| :|
Db 30 RESLR 34

RESULT 10

```

O9QZK5
ID 09QZK5 PRELIMINARY; PRT; 18 AA.
AC 09QZK5;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Desmoplakin (Fragment).
GN
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=C75BL6/J;
RA Green K.J., Guy S.G., Caerhalmi-Friedman P., McLean W.H.I.,
RA Cristiano A.M., Wagner R.M.;
RT "Analysis of the desmoplakin gene reveals striking conservation with
RT other members of the plakin family of cytolinkers."
RL Submitted (May-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF148514; AAF02527.1; -.
DR MGD; MGI:109611; Dsp.
FT NON TER 1 18
SQ SEQUENCE 18 AA; 2289 MW; 3B4IDF23C1E2F960 CRC64;

Query Match
Best Local Similarity 58.1%; Score 18; DB 11; Length 18;
Pred. No. 3.4e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 REXLR 5
DB 2 REXLR 6

RESULT 11
O9QZK4 PRELIMINARY; PRT; 18 AA.
ID 09QZK4;
AC 09QZK4;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Desmoplakin (Fragment).
GN
OS Mus spretus (Western wild mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10096;
RN
RP SEQUENCE FROM N.A.
RA Green K.J., Guy S.G., Caerhalmi-Friedman P., McLean W.H.I.,
RA Cristiano A.M., Wagner R.M.;
RT "Analysis of the desmoplakin gene reveals striking conservation with
RT other members of the plakin family of cytolinkers."
RL Submitted (May-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF148515; AAF02528.1; -.
DR MGD; MGI:109611; Dsp.
FT NON TER 1 18
SQ SEQUENCE 18 AA; 2289 MW; 3B4IDF23C1E2F960 CRC64;

Query Match
Best Local Similarity 58.1%; Score 18; DB 11; Length 18;
Pred. No. 3.4e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 REXLR 5
DB 2 REXLR 6

RESULT 12
O95SN4 PRELIMINARY; PRT; 21 AA.
ID 095SN4

```

```

AC 095SN4;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE GH13116P.
GN BEST:GH10831 OR CG8143.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RA Stapleton M., Broksrein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Fries E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY060690; AAL28238.1; -.
DR FlyBase; FBgn0040502; BEST:GH10831.
SQ SEQUENCE 21 AA; 2463 MW; 6594EB26A78F025 CRC64;

Query Match
Best Local Similarity 58.1%; Score 18; DB 5; Length 21;
Pred. No. 4e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 REXLR 5
DB 4 REXLR 8

RESULT 13
O83274 PRELIMINARY; PRT; 26 AA.
ID 083274;
AC 083274;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE ORF III.
OS cucumber mosaic virus (cucumber mosaic cucumovirus).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Bromoviridae;
OC Cucumovirus.
OX NCBI_TaxID=12305;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=TA3;
RA Masata C., Hayaehi Y., Wang W.Q., Takanami Y.;
RT "Comparison of four satellite RNA isolates of cucumber mosaic virus."
RL Ann. Phytopathol. Soc. Jpn. 56:207-212(1990).
DR EMBL; D10039; BAA00930.1; -.
SQ SEQUENCE 26 AA; 3094 MW; 66B2791D0B5ED89 CRC64;

Query Match
Best Local Similarity 58.1%; Score 18; DB 12; Length 26;
Pred. No. 4.9e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 REXLR 5
DB 18 REXLR 22

RESULT 14
O9J5W8 PRELIMINARY; PRT; 27 AA.
ID 09J5W8;
AC 09J5W8;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Genome polyprotein [Contains: envelope glycoprotein E2 (GP68) (GP70)
DE (NS1)] (Fragment).
OS Hepatitis C virus.

```

```
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20219428; PubMed=10756048;
RA Manzin A., Solforosi L., Debiaggi M., Zara F., Tanzi E., Romano L.,
RA Zanetti A.R., Clementi M.;
RT evolution in perinatal infection.";
RL J. Virol. 74:4327-4334(2000).
DR EMBL; AF192441; AAF65652.1; -.
DR InterPro: IPR002531; HCV NS1.
DR Pfam; PF01560; HCV NS1; 1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
FT NON_TER 1
FT NON_TER 27
SQ SEQUENCE 27 AA; 3053 MW; FC584F0131E665CF CRC64;

Query Match 58.1%; Score 18; DB 12; Length 27;
Best Local Similarity 30.0%; Pred. No. 5.1e+02;
Matches 3; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 REXLRXXXXY 10
Db |||:|
7 REAVRHTTSP 16

RESULT 15
Q9S871 PRELIMINARY; PRT; 30 AA.
AC Q9S871;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE 3.7 kDa napin-like protein small chain (Fragment).
OS Momordica charantia (Bitter melon) (Balsam pear).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Cucurbitales; Cucurbitaceae; Momordica.
OX NCBI_TaxID=3673;
RN [1]
RP SEQUENCE.
RX MEDLINE=97135090; PubMed=8980648;
RA Neumann G.M., Condron R., Polya G.M.;
RT "Purification and sequencing of napin-like protein small and large
RT chains from Momordica charantia and Ricinus communis seeds and
RT determination of sites phosphorylated by plant Ca(2+)-dependent
RT protein kinase.";
RL Biochim. Biophys. Acta 1298:223-240(1996).
SQ SEQUENCE 30 AA; 3737 MW; 21B6BE02D63DECD9 CRC64;

Query Match 58.1%; Score 18; DB 10; Length 30;
Best Local Similarity 80.0%; Pred. No. 5.6e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 REXLR 5
Db |||
13 REQRLR 17

Search completed: December 19, 2002, 16:21:51
Job time : 18.6471 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 19, 2002, 16:16:02 ; Search time 8.67647 Seconds
(without alignments)
110.799 Million cell updates/sec

Title: US-08-653-294C-38

Perfect score: 31

Sequence: 1 REXLRXXXXY 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 15911

Minimum DB seq length: 0

Maximum DB seq length: 60

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 73:*

1: piri:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	67.7	51	2	B31946
2	19	61.3	35	2	D69330
3	19	61.3	57	2	AH2385
4	19	61.3	60	2	S01800
5	18	58.1	33	2	B44906
6	18	58.1	33	2	C97406
7	18	58.1	37	2	AH0844
8	18	58.1	39	2	I56173
9	18	58.1	48	2	S13571
10	18	58.1	50	2	G84198
11	18	58.1	50	2	E81927
12	18	58.1	54	4	JE0018
13	18	58.1	56	2	I56605
14	18	58.1	57	2	A69385
15	18	58.1	57	2	G84072
16	18	58.1	60	2	B75147
17	18	58.1	60	4	JA0077
18	17	54.8	25	2	B85928
19	17	54.8	33	2	C47214
20	17	54.8	47	2	T29970
21	17	54.8	57	2	AB0724
22	17	54.8	60	2	B61510
23	16	51.6	20	2	T01691
24	16	51.6	32	2	A44906
25	16	51.6	38	2	G33646
26	16	51.6	38	2	G70236
27	16	51.6	39	2	A70106
28	16	51.6	46	2	T50847
29	16	51.6	46	2	AH2518

30 16 51.6 49 2 AG3362 hypothetical prote
31 16 51.6 50 2 C82773 hypothetical prote
32 16 51.6 51 2 E82981 50S ribosomal prot
33 16 51.6 51 2 S76541 hypothetical prote
34 16 51.6 51 2 S42357 probable zinc fing
35 16 51.6 51 2 F81545 hypothetical prote
36 16 51.6 51 2 AB2484 hypothetical prote
37 16 51.6 54 2 A26463 ribosomal protein
38 16 51.6 55 2 E82347 ribosomal protein
39 16 51.6 56 2 D64104 ribosomal protein
40 16 51.6 57 2 S07344 nef protein - huma
41 16 51.6 57 2 H95181 hypothetical prote
42 16 51.6 58 2 A18119 hypothetical prote
43 16 51.6 58 2 AH2103 hypothetical prote
44 16 51.6 59 2 D83672 hypothetical prote
45 16 51.6 60 2 S50004 ribosomal protein

ALIGNMENTS

RESULT 1

B31946

hypothetical protein (xhd 5' region) - fruit fly (*Drosophila pseudoobscura*) (fragment)

C:Species: *Drosophila pseudoobscura*

C>Date: 22-Nov-1989 #sequence_revision 22-Nov-1989 #text_change 20-Mar-1998

C:Accession: B31946

R:Riley, M.A.

Mol. Biol. Evol. 6, 33-52, 1989

A:Title: Nucleotide sequence of the Xhd region in *Drosophila pseudoobscura* and an analys

A:Reference number: A31946; MUID:89158785; PMID:2493563

A:Accession: B31946

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-51 <RIL>

A:Cross-references: GB:M33977; NID:gl58807; PID:gl58808

C:Genetics:

A:Gene: FlyBase:Dpae/ry

A:Cross-references: FlyBase:PBgn0012736

Query Match 67.7%; Score 21; DB 2; Length 51;

Best Local Similarity 40.0%; Fred. NO. 43;

Matches 4; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 REXLRXXXXY 10

Db 28 REAMRRQALY 37

RESULT 2

D69330

hypothetical protein AF0644 - *Archaeoglobus fulgidus*

C:Species: *Archaeoglobus fulgidus*

C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999

C:Accession: D69330

R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson

; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.

Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.

Nature 390, 364-370, 1997

A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artach, P.; Kaine, B.P.; Sykes, S.

Smith, H.O.; Woese, C.R.; Venter, J.C.

A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae

A:Reference number: A69250; MUID:98049343; PMID:9389475

A:Accession: D69330

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-35 <KLE>

A:Cross-references: GB:AB001060; GB:AE000782; NID:G2689383; PID:AA090607.1; PID:G2664998

Query Match 61.3%; Score 19; DB 2; Length 35;

Best Local Similarity 80.0%; Fred. NO. 1.le+02;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 REXLR 5
|||
DB 29 REXLR 33

RESULT 3

hypothetical protein asl4640 [imported] - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp.
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
C:Accession: AH2385
R:Kaneko, T.; Nakamura, Y.; Molk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriyuchi, Nakazaki, N.; Shimpou, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena PCC 7120
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AH2385
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-57 <KUR>
A:Cross-references: GB:BA000019; PIDN:BA076339.1; PID:g17133777; GSPDB:GN00179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: asl4640

Query Match 61.3%; Score 19; DB 2; Length 57;
Best Local Similarity 80.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 REXLR 5
|||
DB 22 REXLR 26

RESULT 4

pepsin (EC 3.4.23.-) 3 precursor - North Pacific bluefin tuna (fragment)

C:Species: Thunnus thynnus orientalis (North Pacific bluefin tuna)
C:Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 08-Nov-1996
C:Accession: S01800
R:Tanji, M.; Kageyama, T.; Takahashi, K.
Eur. J. Biochem. 177, 251-259, 1988
A:Title: Tuna pepsinogens and pepsins. Purification, characterization and amino-terminal
A:Reference number: S01798; MUID:89052692; PMID:3142769
A:Accession: S01800
A:Molecule type: protein
A:Residues: 1-60 <TRAN>
C:Superfamily: pepsin
C:Keywords: aspartic proteinase, hydrolase, protein digestion
F:1-35/Domain: propeptide #status experimental <PRO>
F:36-60/Product: pepsin (fragment) #status experimental <MAT>

Query Match 61.3%; Score 19; DB 2; Length 60;
Best Local Similarity 80.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 REXLR 5
|||
DB 12 REXLR 16

RESULT 5

LI protein - human papillomavirus type Xd (fragment)

C:Species: human papillomavirus type Xd
C:Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 26-Aug-1999
C:Accession: B44906
R:van den Brule, A.J.; Snijders, P.J.; Raaphorst, P.M.; Schrijnemakers, H.F.; Delius, H.
J. Clin. Microbiol. 30, 1716-1721, 1992
A:Title: General primer polymerase chain reaction in combination with sequence analysis
A:Reference number: A44889; MUID:92332706; PMID:1321168
A:Accession: B44906

A:Molecule type: DNA
A:Residues: 1-33 <VAN>
A:Cross-references: GB:S40280; NID:g251702; PIDN:AA022572.1; PID:g251703
A:Experimental source: type Xd, cervical carcinoma in situ
A:Note: sequence extracted from NCBI backbone (NCBI:109401, NCBI:109450)
C:Superfamily: papillomavirus LI protein

Query Match 58.1%; Score 18; DB 2; Length 33;
Best Local Similarity 80.0%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 REXLR 5
|||
DB 25 REXLR 29

RESULT 6

hypothetical protein AGR_C_681 [imported] - Agrobacterium tumefaciens (strain C58, Cerec

C:Species: Agrobacterium tumefaciens
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 11-Jan-2002
C:Accession: C97406
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens
A:Reference number: A97359; PMID:11743194
A:Accession: C97406
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-33 <KUR>
A:Cross-references: GB:AE007869; PIDN:AAK86204.1; PID:g15155303; GSPDB:GN00169
C:Genetics:
A:Gene: AGR_C_681
A:Map position: circular chromosome

Query Match 58.1%; Score 18; DB 2; Length 33;
Best Local Similarity 80.0%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 REXLR 5
|||
DB 12 REXLR 16

RESULT 7

hypothetical protein STY2958 [imported] - Salmonella enterica subsp. enterica serovar Typh

C:Species: Salmonella enterica subsp. enterica serovar Typh
A:Note: this species has also been called Salmonella typhi
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 09-Nov-2001
C:Accession: AH0844
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Main, J.; Churcher, Th. T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P.
Nature 413, 846-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A:Reference number: AB0502; PMID:11677608
A:Accession: AH0844
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-37 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD05943.1; PID:g16503914; GSPDB:GN00176
C:Genetics:
A:Gene: STY2958

Query Match 58.1%; Score 18; DB 2; Length 37;
Best Local Similarity 80.0%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 REXLR 5
|||

Db 30 REVLRL 34

RESULT 8

I56173
EBV/C3d receptor - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 21-Jul-2000
C:Accession: I56173
R:Singha, S.K.; Todd, S.C.; Hedrick, J.A.; Speiser, C.L.; Lambiris, J.D.; Tsoukas, C.D.
J. Immunol. 150, 5311-5320, 1993
A:Title: Characterization of the EBV/C3d receptor on the human Jurkat T cell line: evidence for a novel protein.
A:Reference number: I56173; MUID:93294286; PMID:8390533
A:Accession: I56173
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-39 <RES>
A:Cross-references: GB:S62696; NID:G386263; PIDN:AAB27186.1; PID:G386264

Query Match 58.1%; Score 18; DB 2; Length 39;
Best Local Similarity 80.0%; Pred. No. 2.4e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 REXLR 5

|||

Db 23 REFLR 27

RESULT 9

I51371
Hypothetical protein 2 (insertion sequence 1076) - Lactococcus lactis subsp. lactis
C:Species: Lactococcus lactis subsp. lactis
C>Date: 19-Mar-1997 #sequence_revision 06-Jun-1997 #text_change 15-Oct-1999
C:Accession: I51371
R:Huang, D.C.; Novel, M.; Novel, G.
FEMS Microbiol. Lett. 77, 101-106, 1991
A:Title: A transposon-like element on the lactose plasmid of Lactococcus lactis subsp. I
A:Reference number: I513568
A:Accession: I51371
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-48 <HUA>
A:Cross-references: EMBL:X53013; NID:G44078; PIDN:CAA37196.1; PID:G44082
A:Experimental source: insertion sequence 1076; plasmid pUCL22; strain Z270
C:Genetics:
A:Mobile element: insertion sequence 1076

Query Match 58.1%; Score 18; DB 2; Length 48;
Best Local Similarity 40.0%; Pred. No. 2.9e+02;
Matches 4; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 REXLRXXXXY 10

|:|

Db 6 RQELRLKKY 15

RESULT 10

G84198
Hypothetical protein Vng0402h [imported] - Halobacterium sp. NRC-1
C:Species: Halobacterium sp. NRC-1
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: G84198
R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Leitchauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A:Title: Genome sequence of Halobacterium species NRC-1.
A:Reference number: A84160; MUID:20504483; PMID:11016950
A:Accession: G84198
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-50 <STO>

A:Cross-references: GB:AE004437; NID:G10580015; PIDN:AAG18955.1; GSPDB:GN00138
C:Genetics:
A:Gene: VNG0402H

Query Match 58.1%; Score 18; DB 2; Length 50;
Best Local Similarity 80.0%; Pred. No. 3e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 REXLR 5

|||

Db 16 REQLR 20

RESULT 11

E81927
Hypothetical protein NMA0824 [imported] - Neisseria meningitidis (strain Z2491 serogroup C:Species: Neisseria meningitidis
C>Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
C:Accession: E81927
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
A:Reference number: A81775; MUID:20222556; PMID:10761919
A:Accession: E81927
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-50 <PAR>
A:Cross-references: GB:AL162754; GB:AL157959; NID:G7379424; PIDN:CAB84106.1; PID:G737954
A:Experimental source: serogroup A, strain Z2491
C:Genetics:
A:Gene: NMA0822; NMA0824

Query Match 58.1%; Score 18; DB 2; Length 50;
Best Local Similarity 80.0%; Pred. No. 3e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 REXLR 5

|||

Db 42 REELR 46

RESULT 12

JE0018
Probable 6.1K protein pseudogene - cucumber mosaic virus (strain Indonesia, isolates X2)
C:Species: cucumber mosaic virus, CMV
C>Date: 31-Dec-1991 #sequence_revision 28-May-1998 #text_change 20-Oct-2000
C:Accession: JE0018
R:Kaper, J.M.; Tounignant, M.E.; Steen, M.T.
Virology 163, 284-292, 1988
A:Title: Cucumber mosaic virus-associated RNA 5. XI. Comparison of 14 CARNA 5 variants 1
A:Reference number: JE0018; MUID:88179532; PMID:3354198
A:Accession: JE0018
A:Status: translation not shown; conceptual translation of pseudogene
A:Molecule type: genomic RNA
A:Residues: 1-54 <KAP>
A:Cross-references: GB:M20354; GB:M20355; NID:G331670; NID:G331674; PIDN:AAA46396.1; PII
C:Genetics:
A:Map position: segment 5
C:Keywords: pseudogene

Query Match 58.1%; Score 18; DB 2; Length 54;
Best Local Similarity 80.0%; Pred. No. 3.3e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 REXLR 5

|||

Db 46 REELR 50

RESULT 13

I56605
17 alpha-hydroxylase-17,20-desmolase - rat (fragment)

C:Species: Rattus norvegicus (Norway rat)
C:Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 11-Jun-1999
C:Accession: 156605
R:Bučko, E.; Koh, Y.C.; Miyagawa, Y.; Dufau, M.L.
J. Steroid Biochem. Mol. Biol. 52, 209-218, 1995
A:Title: The rat 17 alpha-hydroxylase-17/20-desmolase (CYP17) active site: computerized
A:Reference number: 156605; MUID:95210123; PMID:7696141
A:Accession: 156605
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-56 <RES>
A:Cross-references: EMBL:U14953; NID:g602419; PIDN:AAA85653.1; PID:g602420
C:Genetics:
A:Gene: CYP17
C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
C:Keywords: heme; transmembrane protein

Query Match 58.1%; Score 18; DB 2; Length 56;
Best Local Similarity 80.0%; Pred. No. 3.4e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 REXLR 5
|||
Db 35 REVL 39

RESULT 14
A69385
hypothetical protein AF1081 - Archaeoglobus fulgidus
C:Species: Archaeoglobus fulgidus
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
C:Accession: A69385
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Klinkner, E.F.
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A:Authors: Overbeek, T.; Cotton, M.D.; Spriggs, T.; Artlich, P.; Kaine, B.P.; Sykes, S.
Smith, H.O.; Woese, C.R.; Venter, J.C.
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
A:Reference number: A69385; MUID:98049343; PMID:9389475
A:Accession: A69385
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-57 <KLE>
A:Cross-references: GB:AE001028; GB:AE000782; NID:g2689351; PIDN:AAB90166.1; PID:g264951

Query Match 58.1%; Score 18; DB 2; Length 57;
Best Local Similarity 80.0%; Pred. No. 3.5e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 REXLR 5
|||
Db 11 REVL 15

RESULT 15
G84072
hypothetical protein BH3383 [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: G84072
R:Takami, H.; Nakasone, K.; Takaki, Y.; Mieno, G.; Sasaki, R.; Maui, N.; Fuji, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: G84072
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-57 <STO>
A:Cross-references: GB:AP001518; GB:BA000004; NID:g10175792; PIDN:BA807102.1; GSPDB:GN0
A:Experimental source: strain C-125
C:Genetics:
A:Gene: BH3383

Query Match 58.1%; Score 18; DB 2; Length 57;
Best Local Similarity 80.0%; Pred. No. 3.5e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 REXLR 5
|||
Db 37 REVL 41

Search completed: December 19, 2002, 16:22:57
Job time : 9.67647 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 19, 2002, 16:03:12 ; Search time 4.55882 Seconds
(without alignments)
90.980 Million cell updates/sec

Title: US-08-653-294C-38

Perfect score: 31

Sequence: 1 REXLRXXXXY 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 5116

Minimum DB seq length: 0

Maximum DB seq length: 60

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	19	61.3	35	Y644 ARCFU	O29613 archaeoglob
2	19	61.3	60	PEP3 THUTO	P20141 thunnus thy
3	18	58.1	15	UC27 MAIZE	P80633 zea mays (m
4	18	58.1	39	NPF MONEZ	P41967 moniezia ex
5	18	58.1	57	YAB1 ARCFU	O29182 archaeoglob
6	17	54.8	42	NCAP NDVB	P09459 newcastle d
7	17	54.8	57	GNS SALTY	O8XGE3 salmonella
8	17	54.8	57	YOR3 TTV1	P19278 thermoprote
9	16	51.6	35	NEF HV1H3	P05854 human immun
10	16	51.6	39	Y049 BORBU	O51078 borrelia bu
11	16	51.6	50	POR GLUOX	P80354 gluconobact
12	16	51.6	51	RL33 PSEAE	O9Htn9 pseudomonas
13	16	51.6	51	SPE SEPOF	P80304 speia offic
14	16	51.6	54	RELX SQUAC	P11953 squalus aca
15	16	51.6	55	RL33 VIBCH	O9Kvc7 vibrio chol
16	16	51.6	56	RL33 HAEIN	P44369 haemophilus
17	16	51.6	60	RL30 STRCO	P46789 streptomyce
18	15	48.4	30	CALM LVTPI	P05935 lytechinus
19	15	48.4	39	NPF HELAS	P41321 helix asper
20	15	48.4	43	VPU HV1RH	P12518 human immun
21	15	48.4	43	ARAA KUEPN	Q48433 klebsiella
22	15	48.4	43	Y4KR RHISN	P55536 rhizobium s
23	15	48.4	54	RELX BALAC	P11184 balaenopter
24	15	48.4	54	RELX BALED	P11185 balaenopter
25	15	48.4	59	YH13 ARCFU	O28560 archaeoglob
26	14	45.2	27	SECR CANFA	P09910 canis famli
27	14	45.2	28	ICPP VIPLP	P82475 vipera lebe
28	14	45.2	32	COA2 BPFD	P03677 bacterioph
29	14	45.2	32	GHR4 RAT	P33581 rattus norv
30	14	45.2	32	RIP2 PHYDI	P34967 phytolacca
31	14	45.2	33	MBPI MAIZE	P28794 zea mays (m
32	14	45.2	33	Y50A MYCTU	Q9C556 mycobacteri
33	14	45.2	38	VPU HV1A2	P05949 human immun

34 14 45.2 40 1 YG65 HAEIN P44283 haemophilus
35 14 45.2 43 1 VPU HV1C4 P08803 human immun
36 14 45.2 45 1 MONA DIOCU P02881 dioscoreoph
37 14 45.2 45 1 OSTC LEPMA P28317 leptomis mac
38 14 45.2 45 1 OSTC SPAAU P40148 sparus aua
39 14 45.2 46 1 DIUH PERAM P41538 periplaneta
40 14 45.2 46 1 DIUH ZOONE P82707 zoeteromopsi
41 14 45.2 46 1 LAB9 LACAC Q48501 lactobacill
42 14 45.2 47 1 OSTC XIPGL P02823 xiphias gla
43 14 45.2 48 1 RL34 MYCGL P47704 mycoplasma
44 14 45.2 50 1 YCX7 OBHNO Q9mcm9 oenothera h
45 14 45.2 52 1 PPLA_CHICK P26677 gallus gall

ALIGNMENTS

RESULT 1

Y644 ARCFU
ID Y644 ARCFU STANDARD; PRT; 35 AA.
AC O29613;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein AF0644.
GN AF0644.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC Archaeoglobaceae; Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Kyrpides N.C.,
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyripides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.P., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
reducing archaeon Archaeoglobus fulgidus.";
RL Nature 390:364-370(1997).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

DR EMBL: AE001060; AAB90607.1; --
DR TIGR: AF0644; --
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 35 AA; 4056 MW; 34FB802042F10226 CRC64;

Query Match 61.3%; Score 19; DB 1; Length 35;
Best Local Similarity 80.0%; Pred. No. 45;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 REXLR 5
|||
Db 29 RETLR 33

RESULT 2
PEP3_THUTO

```

ID  PEPT THUTO      STANDARD;      PRT;      60 AA.
AC  P20141.
DT  01-FEB-1991 (Rel. 17, Created)
DT  01-FEB-1991 (Rel. 17, Last sequence update)
DT  15-JUN-2002 (Rel. 41, Last annotation update)
DE  Pepsin 3 precursor (EC 3.4.23.-) (Fragment).
OS  Thunnus thynnus orientalis (North Pacific bluefin tuna).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC  Acanthomorpha; Acanthopterygii; Perciformes; Scombroidei;
OC  Scombridae; Thunnus.
OX  NCBI_TaxID=8238;
RN  [1]
RP  SEQUENCE.
RX  MEDLINE=89052692; PubMed=3142769;
RA  Tanji M., Kageyama T., Takahashi K.;
RT  "Tuna pepsinogens and pepsins. Purification, characterization and
RT  amino-terminal sequences."
RL  Eur. J. Biochem. 177:251-259(1988).
CC  -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.
DR  PIR; S01800; S01800.
DR  HSSP; P20142; 1HTR.
DR  MEROPS; A01.001; -.
DR  InterPro: IPR001461; AsproteaseA1.
DR  InterPro: IPR001969; Asprotease_site.
DR  Pfam; PF00026; asp; 1.
DR  PROSITE; PS00141; ASP_PROTEASE; PARTIAL.
KW  Hydrolyase; Aspartyl protease; Zymogen.
FT  PROPEP 1 35 ACTIVATION PEPTIDE.
FT  CHAIN 36 >60 PEPsin 3.
FT  NON TER 60 60
SQ  SEQUENCE 60 AA; 6970 MW; F9E6559B2728A892 CRC64;

Query Match      61.3%; Score 19; DB 1; Length 60;
Best Local Similarity 80.0%; Pred. No. 80;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY  1 REXLR 5
    |||
Db  12 RESLR 16

RESULT 3
UC27_MAIZE      STANDARD;      PRT;      15 AA.
AC  P80633.
DT  01-OCT-1996 (Rel. 34, Created)
DT  01-OCT-1996 (Rel. 34, Last sequence update)
DT  15-JUL-1999 (Rel. 38, Last annotation update)
DE  Unknown protein from 2D-page of etiolated coleoptile (Spot 688)
DE  (Fragment).
OS  Zea mays (Maize).
OC  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC  Panicoideae; Andropogoneae; Zea.
OX  NCBI_TaxID=4577;
RN  [1]
RP  SEQUENCE.
RC  TISSUE=Coleoptile;
RA  Touzet P., Ricciardi F., Morin C., Damerval C., Huet J.-C.,
RA  Penollet J.-C., Zivy M., de Vienne D.;
RT  "The maize two dimensional gel protein database: towards an integrated
RT  genome analysis program."
RL  Theor. Appl. Genet. 93:997-1005(1996).
CC  -1- MISCELLANEOUS: ON THE 2D-GEL, THE DETERMINED PI OF THIS UNKNOWN
CC  PROTEIN IS: 6.4. ITS MW IS: 48.4 KDa.
DE  -1- SIMILARITY: TO XENOPUS HISTONIC BINDING PROTEIN N1/N2 AND RABBIT
CC  AND HUMAN NUCLEAR AUTOMITOTIC SPERM PROTEIN.
DR  Maize-2DPAGE; P80633; COLEOPTILE.
DR  MaizeDB; 123958; -.
FT  NON TER 1 1
FT  NON TER 15 15
SQ  SEQUENCE 15 AA; 1853 MW; CA0E12A5DADBD8C7 CRC64;

```

```

Query Match      58.1%; Score 18; DB 1; Length 15;
Best Local Similarity 40.0%; Pred. No. 35;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY  1 REXLR 5
    |||
Db  2 REQLR 6

RESULT 4
NPF_MONEX      STANDARD;      PRT;      39 AA.
AC  P41967.
DT  01-NOV-1995 (Rel. 32, Created)
DT  01-NOV-1995 (Rel. 32, Last sequence update)
DT  16-OCT-2001 (Rel. 40, Last annotation update)
DE  Neuropeptide F (NPF).
OS  Montezia expansa (Sheep tapeworm).
OC  Eukaryota; Metazoa; Platyhelminthes; Cestoda; Eucestoda;
OC  Cyclophyllidae; Anoplocephalidae; Montezia.
OX  NCBI_TaxID=28841;
RN  [1]
RP  SEQUENCE.
RA  Maule A.G., Shaw C., Halton D.W., Thim L., Johnston C.F.,
RA  Fairweather I., Buchanan K.D.;
RT  "Neuropeptide F: a novel parasitic flatworm regulatory peptide from
RT  Montezia expansa (Cestoda: Cyclophyllidae)."
RL  Parasitology 102:309-316(1991).
RN  [2]
RP  CHARACTERIZATION.
RX  MEDLINE=93095525; PubMed=1461689;
RA  Maule A.G., Shaw C., Halton D.W., Brennan G.P., Johnston C.F.,
RA  Moore S.;
RT  "Neuropeptide F (Montezia expansa): localization and characterization
RT  using specific antisera."
RL  Parasitology 105:505-512(1992).
CC  -1- FUNCTION: MAY HAVE AN IMPORTANT PHYSIOLOGICAL ROLE IN
CC  NEUROREGULATION.
CC  -1- SUBCELLULAR LOCATION: Secreted.
CC  -1- SIMILARITY: BELONGS TO THE NPY / PPY / PYY FAMILY.
DR  InterPro: IPR001955; Pancreatic_hormn.
DR  Pfam; PF00159; hormone; 1.
DR  SMART; SM00309; PAH; 1.
DR  PROSITE; PS00265; PANCREATIC_HORMONE_1; 1.
DR  PROSITE; PS50276; PANCREATIC_HORMONE_2; 1.
KW  Neuropeptide; Amidation.
FT  MOD RRS 39 39 AMIDATION.
SQ  SEQUENCE 39 AA; 4594 MW; 2D61A76927DEA732 CRC64;

Query Match      58.1%; Score 18; DB 1; Length 39;
Best Local Similarity 40.0%; Pred. No. 95;
Matches 4; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY  1 REXLRXXXXY 10
    |||
Db  21 RDLRLQINEY 30

RESULT 5
YA81_ARCFU      STANDARD;      PRT;      57 AA.
AC  O29182.
DT  16-OCT-2001 (Rel. 40, Created)
DT  16-OCT-2001 (Rel. 40, Last sequence update)
DT  16-OCT-2001 (Rel. 40, Last annotation update)
DE  Hypothetical protein AF1081.
GN  AF1081.
OS  Archaeoglobus fulgidus.
OC  Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC  Archaeoglobaceae; Archaeoglobus.
OX  NCBI_TaxID=2234;
RN  [1]

```

RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., Hickey E.K., Peterson K.E.,
RA Ketchum K.A., Dodson R.J., Winn M., Graham D.E., Kyprides N.C.,
RA Richardson D.L., Kerlavage A.R., Lee N.H., Sutton G.G., Gill S.,
RA Fleischmann R.D., Quackenbush J., McKenney K., Adams M.D., Loftus B.,
RA Kirkness E.F., Dougherty B.A., Weidman J.F., McDonald L., Utterback T.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artach P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon *Archaeoglobus fulgidus*.";
RL Nature 390:364-370(1997).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE001028; AAB90166.1; -;
DR TIGR; AF1081; -;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 57 AA; 6698 MW; 7CA662578176B22 CRC64;

Query Match 58.1%; Score 18; DB 1; Length 57;
Best Local Similarity 80.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 REVL 5
DB 11 REVL 15

RESULT 6
NCAP_NDVB STANDARD; PRT; 42 AA.
AC P09459;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 01-MAR-1992 (Rel. 21, Last annotation update)
DE Nucleocapsid protein (Fragment).
GN NP.
OS Newcastle disease virus (strain Beaudette C/45) (NDV).
OC Viruses; SERNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Rubulavirus.
OX NCBI_TaxID=11178;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85274370; PubMed=4024452;
RA Kurilla M.G., Stone H.O., Keene J.D.;
RT "RNA sequence and transcriptional properties of the 3' end of the
RT Newcastle disease virus genome.";
RL Virology 145:203-212(1985).
CC -!- FUNCTION: MOST ABUNDANT PROTEIN IN THE VIRION AND AN IMPORTANT
CC ELEMENT INTERACTING WITH M PROTEIN DURING VIRION FORMATION.
CC WELL AS INTERACTING WITH M PROTEIN DURING VIRION FORMATION.
CC -!- SIMILARITY: BELONGS TO THE PARAMYXOVIRUSES NUCLEOCAPSID FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; M11204; AAA46676.1; -;
KW Nucleocapsid.
FT NON_TER 42 42
SQ SEQUENCE 42 AA; 4917 MW; 58946F273B21DBB0 CRC64;

Query Match 54.8%; Score 17; DB 1; Length 42;
Best Local Similarity 40.0%; Pred. No. 2e+02;
Matches 4; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 REVLXXXXX 10
DB 26 REVLQSRPY 35

RESULT 7
GNS_SALTY STANDARD; PRT; 57 AA.
AC Q8XG53;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Gns protein.
GN GNS OR STM1809 OR STY1938.
OS Salmonella typhimurium, and
OS Salmonella typhi.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602, 601;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhimurium; STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du P., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of *Salmonella enterica* serovar Typhimurium
RT LT2.";
RL Nature 413:852-856(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhi; STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Bentley S.D., Holden M.T.G., Sebaihia M.,
RA Churcher C., Mungall K.L., Bentley S.D., Chillingworth T., Connor P.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrall B.G.;
RT "Complete genome sequence of a multiple drug resistant *Salmonella*
RT enterica serovar Typhi CT18.";
RL Nature 413:848-852(2001).
CC -!- SIMILARITY: BELONGS TO THE GNS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE008780; AAL20724.1; -;
DR EMBL; AL627272; CAD05492.1; -;
DR StyGene; SG77777; gns.
KW Complete proteome.
SQ SEQUENCE 57 AA; 6482 MW; 45C7F4A9CC937A9 CRC64;

Query Match 54.8%; Score 17; DB 1; Length 57;
Best Local Similarity 30.0%; Pred. No. 2.7e+02;

Matches 3; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 1 REXLXXXXX 10
 Db 41 RETMKLGEY 50

RESULT 8
 YOR3_TTV1 STANDARD; PRT; 57 AA.
 ID YOR3_TTV1
 AC P19278;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 01-NOV-1990 (Rel. 16, Last annotation update)
 DB Hypothetical 6.9 kDa protein.
 OS Thermoproteus tenax virus 1 (strain KRA1) (TTV1).
 OC Viruses; dsDNA viruses, no RNA stage; Lipothirixviridae;
 CC Lipothirixvirus.
 CC NCBI_TaxId=10480;
 CC [1]
 RN SEQUENCE FROM N.A.
 RA Neumann H.;
 RL Submitted (MAR-1989) to the EMBL/Genbank/DBJ databases.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; X14855; CAA32971.1; -.
 DR Hypothetical protein.
 KW SEQUENCE 57 AA; 6914 MW; B907370704263763 CRC64;
 SQ

Query Match 54.8%; Score 17; DB 1; Length 57;
 Best Local Similarity 44.4%; Pred. No. 2.7e+02;
 Matches 4; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 2 EXLXXXXXX 10
 Db 28 ETLRIDRKY 36

RESULT 9
 NEF_HV1H3 STANDARD; PRT; 35 AA.
 ID NEF_HV1H3
 AC P05854;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Negative factor (F-protein) (27 kDa protein) (3'ORF) (Fragment).
 GN NEF.
 OS Human immunodeficiency virus type 1 (HXB3 isolate) (HIV-1).
 CC Viruses; Retroviridae; Retroviridae; Lentivirus.
 CC NCBI_TaxId=11707;
 CC [1]
 RN SEQUENCE FROM N.A.
 RA MEDLINE=85228248; PubMed=2988795;
 RA Crowl R., Ganguly K., Gordon M., Conroy R., Schaber M., Kramer R.,
 RA Shaw G.M., Wong-Staal F., Reddy E.P.;
 RA "HIV-III env gene products synthesized in E. coli are recognized by
 RA antibodies present in the sera of AIDS patients.";
 RA Cell 41:979-986 (1985).
 RN [2]
 RP POST-TRANSLATIONAL MODIFICATIONS, AND FUNCTION.
 RX MEDLINE=86039140; PubMed=3118220;
 RX Guy B., Kieny M.-P., Riviere Y., le Peuch C., Dotti K., Girard M.,
 RA Montagnier L., Lecocq J.-P.;
 RA "HIV F/3' orf encodes a phosphorylated GTP-binding protein resembling
 RT an oncogene product.";
 RL Nature 330:266-269 (1987).

CC -I- FUNCTION: NEF has GTPase, GTP-binding and autophosphorylating
 CC activities. It seems to down-regulate the CD4(T4) antigen.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; M14100; AAA44680.1; -.
 DR HIV; M14100; NEFSHXB3.
 DR InterPro; IPR001558; HIV_Nef.
 DR Pfam; PF00469; P-protein; 1.
 DR ProDom; PD000031; HIV_Nef; 1.
 DR AIDS; Myristate; GTP-binding.
 FT LIPID 2
 FT NON_TER 35 35 MYRISTATE.
 SQ SEQUENCE 35 AA; 3714 MW; F66354APBEIDFCFE CRC64;
 SQ

Query Match 51.6%; Score 16; DB 1; Length 35;
 Best Local Similarity 60.0%; Pred. No. 3e+02;
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 REXLR 5
 Db 17 RERMR 21

RESULT 10
 Y049_BORBU STANDARD; PRT; 39 AA.
 ID Y049_BORBU
 AC O51078;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein BB0049.
 GN BB0049.
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OC Bacteria; Spirochaetales; Spirochaetales; Borrelia.
 CC NCBI_TaxId=139;
 CC [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=ATCC 35210 / B31;
 RX MEDLINE=98065943; PubMed=9403685;
 RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
 RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
 RA Dougherty B., Tomb J.F., Fleischmann R.D., Richardson D.,
 RA Peterson J., Keitelavage A.R., Quackenbush J., Salzberg S., Hanson M.,
 RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
 RA Uterback T., Matthey L., McDonald L., Artiach P., Bowman C.,
 RA Garland S., Fujii C., Cotton M.D., Horat K., Roberts K., Hatch B.,
 RA Smith H.O., Venter J.C.;
 RA "Genomic sequence of a Lyme disease spirochete, Borrelia
 RT burgdorferi.";
 RL Nature 390:580-586 (1997).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AE001118; AAC66440.1; -.
 DR TIGR; BB0049; -.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 39 AA; 4343 MW; 4DBDE640D709233D CRC64;
 SQ

Query Match 51.6%; Score 16; DB 1; Length 39;
 Best Local Similarity 60.0%; Pred. No. 3.4e+02;


```

RN [1]
RP SEQUENCE.
RC TISSUE=Ovary; PubMed=3780747;
RX MEDLINE=87054035;
RA Ballebach E.E., Gowan L.K., Schwabe C., Steinetz B.G., O'Byrne E.,
RT Callard I.P.,
RT "Isolation, purification, and the sequence of relaxin from spiny
RT dogfish (Squalus acanthias).";
RL Eur. J. Biochem. 161:335-341(1986).
CC -1- FUNCTION: THE FUNCTION OF RELAXIN IN AN OVIPAROUS SPECIES IS NOT
CC YET KNOWN.
CC -1- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
CC DISULFIDE BONDS.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR PIR; A26463; A26463.
DR PIR; B26463; B26463.
DR HSSP; P04090; 6RLX.
DR InterPro; IPR004825; Ins/IGF/relax.
DR SMART; SM00078; IIGF; 1.
DR PROSITE; PS00262; INSULIN; 1.
KW Insulin family; Hormone.
FT CHAIN 1 30 RELAXIN B CHAIN.
FT NON CONS 30 31
FT CHAIN 31 54 RELAXIN A CHAIN.
FT MOD RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT DISULFID 13 41 INTERCHAIN.
FT DISULFID 25 54 INTERCHAIN.
FT DISULFID 40 45
SQ SEQUENCE 54 AA; 5910 MW; 1B7206773AD32A5B CRC64;

Query Match 51.6%; Score 16; DB 1; Length 54;
Best Local Similarity 60.0%; Pred. No. 4.8e+02;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 REXLR 5
DB 15 REFIR 19

RESULT 15
RL33_VIRCH
ID RL33_VIRCH STANDARD; PRT; 55 AA.
AC Q9KVC7;
DT 16-OCT-2001 (Rel. 40; Created)
DT 16-OCT-2001 (Rel. 40; Last sequence update)
DT 16-OCT-2001 (Rel. 40; Last annotation update)
DE 50S ribosomal protein L33.
GN RPLM6 OR VC0219.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=El Tor N16961 / Serotype O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nieman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae";
RL Nature 406:477-483(2000).
CC -1- SIMILARITY: BELONGS TO THE L33P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

```

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE004112; AAF93395.1; -.
DR TIGR; VC0219; -.
DR InterPro; IPR001705; Ribosomal_L33.
DR Pfam; PF00477; Ribosomal_L33; 1.
DR ProDom; PD002595; Ribosomal_L33; 1.
DR TIGRFAMs; TIGR01023; rplm6_bact; 1.
DR PROSITE; PS00582; RIBOSOMAL_L33; FALSE_NEG.
KW Ribosomal protein; Complete proteome.
SQ SEQUENCE 55 AA; 6394 MW; 9D7BB702D663678A CRC64;

Query Match 51.6%; Score 16; DB 1; Length 55;
Best Local Similarity 60.0%; Pred. No. 4.9e+02;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 REXLR 5
DB 6 REKIR 10

```

Search completed: December 19, 2002, 16:19:43
Job time : 5.55882 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 19, 2002, 16:14:32 ; Search time 21.1765 Seconds
(without alignments)
116.760 Million cell updates/sec

Title: US-08-653-294C-36

Perfect score: 58

Sequence: 1 YRLAIRRIALRY 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 45785

Minimum DB seq length: 0

Maximum DB seq length: 60

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mmc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30	51.7	53	Q9PE58	Q9PE58 xylella fas
2	30	51.7	54	O89888	O89888 budgerigar
3	30	51.7	54	O89889	O89889 budgerigar
4	30	51.7	54	O89890	O89890 budgerigar
5	30	51.7	54	O89891	O89891 budgerigar
6	30	51.7	54	O89892	O89892 budgerigar
7	30	51.7	54	O89893	O89893 budgerigar
8	30	51.7	54	O89894	O89894 budgerigar
9	30	51.7	54	O89895	O89895 budgerigar
10	30	51.7	54	O89897	O89897 budgerigar
11	30	51.7	54	O89898	O89898 budgerigar
12	30	51.7	54	O89899	O89899 budgerigar
13	30	51.7	54	O89900	O89900 budgerigar
14	30	51.7	54	O89901	O89901 budgerigar
15	30	51.7	54	O89902	O89902 budgerigar
16	30	51.7	54	O89903	O89903 budgerigar

17	30	51.7	54	12	O89904	O89904 budgerigar
18	30	51.7	54	12	O89896	O89896 budgerigar
19	29.5	50.9	42	11	O54741	O54741 mus musculus
20	28	48.3	35	5	O61237	O61237 onchocerca
21	28	48.3	52	10	O9SK54	O9SK54 arabidopsis
22	28	48.3	52	16	O932A2	O932A2 staphylococ
23	28	48.3	58	16	O9CHV1	O9CHV1 lactococcus
24	28	48.3	59	7	O78094	O78094 homo sapien
25	27	46.6	37	13	O42150	O42150 ambystoma m
26	27	46.6	39	2	Q9R5J2	Q9R5J2 streptococ
27	27	46.6	54	12	O91H00	O91H00 mumps virus
28	27	46.6	57	12	P89006	P89006 mumps virus
29	27	46.6	57	12	P89009	P89009 mumps virus
30	27	46.6	57	12	P89011	P89011 mumps virus
31	27	46.6	57	12	P89014	P89014 mumps virus
32	27	46.6	57	12	P89016	P89016 mumps virus
33	27	46.6	57	12	P89017	P89017 mumps virus
34	27	46.6	57	12	O98352	O98352 mumps virus
35	27	46.6	57	12	Q9WAF2	Q9WAF2 mumps virus
36	27	46.6	57	12	Q9WAGO	Q9WAGO mumps virus
37	27	46.6	57	12	O9WAH4	O9WAH4 mumps virus
38	27	46.6	57	12	O9QEV1	O9QEV1 mumps virus
39	27	46.6	57	12	O9QEV0	O9QEV0 mumps virus
40	27	46.6	57	12	O9QEU8	O9QEU8 mumps virus
41	27	46.6	57	12	O9QEU7	O9QEU7 mumps virus
42	27	46.6	57	12	O9QEU6	O9QEU6 mumps virus
43	27	46.6	57	12	O9QEU5	O9QEU5 mumps virus
44	27	46.6	57	12	O9ICS9	O9ICS9 mumps virus
45	27	46.6	57	12	O91H01	O91H01 mumps virus

ALIGNMENTS

RESULT 1

Q9PE58	PRELIMINARY;	PRT;	53 AA.
ID	O9PE58		
AC	O9PE58;		
DC	01-OCT-2000 (Tremblrel. 15, Created)		
DT	01-OCT-2000 (Tremblrel. 15, Last sequence update)		
DT	01-JUN-2002 (Tremblrel. 21, Last annotation update)		
DE	50S ribosomal protein L30.		
GN	XFL170.		
OS	Xylella fastidiosa.		
OC	Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;		
OC	Xylella.		
OX	NCBI_TaxID=2371;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=9A5C;		
RX	MEDLINE=20365717; PubMed=10910347;		
RA	Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,		
RA	Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,		
RA	Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,		
RA	Bueno M.R.P., Camargo A.A., Carraro L.E.A., Carraro D.M., Carrer H.,		
RA	Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,		
RA	Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,		
RA	Faccinani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,		
RA	Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,		
RA	Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,		
RA	Ho P.L., Hohlseil J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,		
RA	Krieger J.B., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,		
RA	Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,		
RA	Machado M.A., Martins A.M.B.N., Madeira H.M.F., Marino C.L.,		
RA	Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,		
RA	Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,		
RA	Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.B.S.,		
RA	Nhani A.J., Nobrega F.G., Nunes L.R., Oliveira M.A.,		
RA	de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,		
RA	Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pasquero J.B.,		
RA	Quaggio R.B., Roberto P.G., Rodrigues V. de Rosa A.J.M.,		
RA	de Rosa V.B. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,		
RA	da Silva A.C.R., da Silva F.R., da Silva W.A. Jr.,		

```

RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Trufi D., Tsai S.M., Tsubako M.H.,
RA Vallada H., Van Slyks M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.,
RT "The genome sequence of the plant pathogen Xylella fastidiosa.",
RL Nature 406:151-159(2000).
DR EMBL; AF003951; AAF83980.1; -.
DR InterPro; IPR000517; Ribosomal_L30.
DR Pfam; PF00327; Ribosomal_L30; 1.
DR TIGRFAMs; TIGR01308; rplM_dact; 1.
KW Complete proteome.
SQ SEQUENCE 53 AA; 6109 MW; EF7E8A2128BEAD36 CRC64;

Query Match 51.7%; Score 30; DB 16; Length 53;
Best Local Similarity 45.5%; Pred. No. 1.3e+02;
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 YRLAIRRLRLR 11
DB 14 HRLSVRALGLR 24

RESULT 2
089888 PRELIMINARY; PRT; 54 AA.
ID 089888
AC 089888;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Large T and small t antigens (Fragment).
OS Budderigar fledgling disease virus (BFDV).
OC Viruses; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.
OX NCBI_TaxID=10625;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MCFL97;
RA Phalen D.N., Wilson V.G., Gaskin J.M., Derr J.N., Graham D.L.;
RT "Genetic Diversity in 20 Variants of the Avian Polyomavirus.";
RL Submitted (MAR-1998) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF054402; AAC33626.1; -.
DR InterPro; IPR001623; DnaJ_N.
DR Pfam; PF00226; DnaJ_1.
DR SMART; SM00271; DnaJ_1.
DR PROSITE; PS50076; DnaJ_2; 1.
FT NON_TER 54
FT SEQUENCE 54 AA; 6077 MW; SAF094925DE8C997 CRC64;

Query Match 51.7%; Score 30; DB 12; Length 54;
Best Local Similarity 54.5%; Pred. No. 1.4e+02;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 RLAIRRLALRY 12
DB 23 KTAAYRTALKY 33

RESULT 3
089889 PRELIMINARY; PRT; 54 AA.
ID 089889
AC 089889;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Large T and small t antigens (Fragment).
OS Budderigar fledgling disease virus (BFDV).
OC Viruses; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.
OX NCBI_TaxID=10625;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LBM192;
RA Phalen D.N., Wilson V.G., Gaskin J.M., Derr J.N., Graham D.L.;
RT "Genetic Diversity in 20 Variants of the Avian Polyomavirus.";
RL Submitted (MAR-1998) to the EMBL/Genbank/DBJ databases.

```

```

DR EMBL; AF054403; AAC33627.1; -.
DR InterPro; IPR001623; DnaJ_N.
DR Pfam; PF00226; DnaJ_1.
DR SMART; SM00271; DnaJ_1.
DR PROSITE; PS50076; DnaJ_2; 1.
FT NON_TER 54
FT SEQUENCE 54 AA; 6077 MW; SAF094925DE8C997 CRC64;

Query Match 51.7%; Score 30; DB 12; Length 54;
Best Local Similarity 54.5%; Pred. No. 1.4e+02;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 RLAIRRLALRY 12
DB 23 KTAAYRTALKY 33

RESULT 4
089890 PRELIMINARY; PRT; 54 AA.
ID 089890
AC 089890;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Large T and small t antigens (Fragment).
OS Budderigar fledgling disease virus (BFDV).
OC Viruses; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.
OX NCBI_TaxID=10625;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GCAZ92;
RA Phalen D.N., Wilson V.G., Gaskin J.M., Derr J.N., Graham D.L.;
RT "Genetic Diversity in 20 Variants of the Avian Polyomavirus.";
RL Submitted (MAR-1998) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF054404; AAC33628.1; -.
DR InterPro; IPR001623; DnaJ_N.
DR Pfam; PF00226; DnaJ_1.
DR SMART; SM00271; DnaJ_1.
DR PROSITE; PS50076; DnaJ_2; 1.
FT NON_TER 54
FT SEQUENCE 54 AA; 6077 MW; SAF094925DE8C997 CRC64;

Query Match 51.7%; Score 30; DB 12; Length 54;
Best Local Similarity 54.5%; Pred. No. 1.4e+02;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 RLAIRRLALRY 12
DB 23 KTAAYRTALKY 33

RESULT 5
089891 PRELIMINARY; PRT; 54 AA.
ID 089891
AC 089891;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Large T and small t antigens (Fragment).
OS Budderigar fledgling disease virus (BFDV).
OC Viruses; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.
OX NCBI_TaxID=10625;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ECIX91;
RA Phalen D.N., Wilson V.G., Gaskin J.M., Derr J.N., Graham D.L.;
RT "Genetic Diversity in 20 Variants of the Avian Polyomavirus.";
RL Submitted (MAR-1998) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF054405; AAC33629.1; -.
DR InterPro; IPR001623; DnaJ_N.
DR Pfam; PF00226; DnaJ_1.
DR SMART; SM00271; DnaJ_1.
DR PROSITE; PS50076; DnaJ_2; 1.

```

FT NON_TER 54 54
 SQ SEQUENCE 54 AA; 6077 MW; 5AF094925DE8C997 CRC64;
 Query Match 51.7%; Score 30; DB 12; Length 54;
 Best Local Similarity 54.5%; Pred. No. 1.4e+02;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 RLAIIRIALRY 12
 : ||||| :
 Db 23 KTA YRRRTALKY 33

RESULT 6

ID O89892 PRELIMINARY; PRT; 54 AA.
 AC O89892;
 DT 01-NOV-1998 (TREMELrel. 08, Created)
 DT 01-NOV-1998 (TREMELrel. 08, Last sequence update)
 DT 01-JUN-2001 (TREMELrel. 17, Last annotation update)
 DE Large T and small t antigens (Fragment).
 OS Budgerigar fledgling disease virus (BFDV).
 OC Viruses; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.
 OX NCBI_TaxID=10625;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LB85;
 RA Phalen D.N., Wilson V.G., Gaskin J.M., Derr J.N., Graham D.L.;
 RT "Genetic Diversity in 20 Variants of the Avian Polyomavirus";
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF054406; AAC33630.1;
 DR InterPro; IPR001623; DnaJ_N.
 DR Pfam; PF00226; DnaJ; 1.
 DR SMART; SM00271; DnaJ; 1.
 DR PROSITE; PS50076; DnaJ 2; 1.
 FT NON_TER 54 54
 SQ SEQUENCE 54 AA; 6077 MW; 5AF094925DE8C997 CRC64;

Query Match 51.7%; Score 30; DB 12; Length 54;
 Best Local Similarity 54.5%; Pred. No. 1.4e+02;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 RLAIIRIALRY 12
 : ||||| :
 Db 23 KTA YRRRTALKY 33

RESULT 7

ID O89893 PRELIMINARY; PRT; 54 AA.
 AC O89893;
 DT 01-NOV-1998 (TREMELrel. 08, Created)
 DT 01-NOV-1998 (TREMELrel. 08, Last sequence update)
 DT 01-JUN-2001 (TREMELrel. 17, Last annotation update)
 DE Large T and small t antigens (Fragment).
 OS Budgerigar fledgling disease virus (BFDV).
 OC Viruses; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.
 OX NCBI_TaxID=10625;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=YCF187;
 RA Phalen D.N., Wilson V.G., Gaskin J.M., Derr J.N., Graham D.L.;
 RT "Genetic Diversity in 20 Variants of the Avian Polyomavirus";
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF054407; AAC33631.1;
 DR InterPro; IPR001623; DnaJ_N.
 DR Pfam; PF00226; DnaJ; 1.
 DR SMART; SM00271; DnaJ; 1.
 DR PROSITE; PS50076; DnaJ 2; 1.
 FT NON_TER 54 54
 SQ SEQUENCE 54 AA; 6077 MW; 5AF094925DE8C997 CRC64;

Query Match 51.7%; Score 30; DB 12; Length 54;
 Best Local Similarity 54.5%; Pred. No. 1.4e+02;

Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 2 RLAIIRIALRY 12
 : ||||| :
 Db 23 KTA YRRRTALKY 33

RESULT 8

ID O89894 PRELIMINARY; PRT; 54 AA.
 AC O89894;
 DT 01-NOV-1998 (TREMELrel. 08, Created)
 DT 01-NOV-1998 (TREMELrel. 08, Last sequence update)
 DT 01-JUN-2001 (TREMELrel. 17, Last annotation update)
 DE Large T and small t antigens (Fragment).
 OS Budgerigar fledgling disease virus (BFDV).
 OC Viruses; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.
 OX NCBI_TaxID=10625;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SCIL88;
 RA Phalen D.N., Wilson V.G., Gaskin J.M., Derr J.N., Graham D.L.;
 RT "Genetic Diversity in 20 Variants of the Avian Polyomavirus";
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF054408; AAC33632.1;
 DR InterPro; IPR001623; DnaJ_N.
 DR Pfam; PF00226; DnaJ; 1.
 DR SMART; SM00271; DnaJ; 1.
 DR PROSITE; PS50076; DnaJ 2; 1.
 FT NON_TER 54 54
 SQ SEQUENCE 54 AA; 6077 MW; 5AF094925DE8C997 CRC64;

Query Match 51.7%; Score 30; DB 12; Length 54;
 Best Local Similarity 54.5%; Pred. No. 1.4e+02;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 RLAIIRIALRY 12
 : ||||| :
 Db 23 KTA YRRRTALKY 33

RESULT 9

ID O89895 PRELIMINARY; PRT; 54 AA.
 AC O89895;
 DT 01-NOV-1998 (TREMELrel. 08, Created)
 DT 01-NOV-1998 (TREMELrel. 08, Last sequence update)
 DT 01-JUN-2001 (TREMELrel. 17, Last annotation update)
 DE Large T and small t antigens (Fragment).
 OS Budgerigar fledgling disease virus (BFDV).
 OC Viruses; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.
 OX NCBI_TaxID=10625;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BDTX88;
 RA Phalen D.N., Wilson V.G., Gaskin J.M., Derr J.N., Graham D.L.;
 RT "Genetic Diversity in 20 Variants of the Avian Polyomavirus";
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF054409; AAC33633.1;
 DR InterPro; IPR001623; DnaJ_N.
 DR Pfam; PF00226; DnaJ; 1.
 DR SMART; SM00271; DnaJ; 1.
 DR PROSITE; PS50076; DnaJ 2; 1.
 FT NON_TER 54 54
 SQ SEQUENCE 54 AA; 6077 MW; 5AF094925DE8C997 CRC64;

Query Match 51.7%; Score 30; DB 12; Length 54;
 Best Local Similarity 54.5%; Pred. No. 1.4e+02;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 RLAIIRIALRY 12
 : ||||| :
 Db 23 KTA YRRRTALKY 33

RESULT	ID	PRELIMINARY;	PRT;	54 AA.
089897	ID	089897		
AC	089897			
DT	01-NOV-1998	(TREMBlrel. 08, Created)		
DT	01-NOV-1998	(TREMBlrel. 08, Last sequence update)		
DT	01-JUN-2001	(TREMBlrel. 17, Last annotation update)		
DE	Large T and small t antigens (Fragment).			
OS	Budgerigar fledgling disease virus (BFDV).			
OC	Viruses; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.			
OX	NCBI_Taxid=10625;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=BDTX89;			
RA	Phalen D.N., Wilson V.G., Gaskin J.M., Derr J.N., Graham D.L.;			
RT	"Genetic Diversity in 20 Variants of the Avian Polyomavirus. ";			
RL	Submitted (MAR-1998) to the EMBL/Genbank/DBJ databases.			
DR	EMBL; AF054411; AAC93635.1; -			
DR	InterPro; IPR001623; DnaJ_N.			
DR	Pfam; PF00226; DnaJ; 1.			
DR	SMART; SM00271; DnaJ; 1.			
DR	PROSITE; PSS0076; DNaJ_2; 1.			
FT	NON_TER	54	54	
SQ	SEQUENCE	54 AA;	6077 MW;	SAF094925DEBC997 CRC64;
Query Match		51.7%;	Score 30;	DB 12; Length 54;
Best Local Similarity		54.5%;	Pred. No. 1.4e+02;	
Matches	6;	Conservative	2;	Mismatches 3; Indels 0; Gaps 0

```

Oy      2 RLAIRRLALRY 12
       : | | | | :
Db      23 KTAAYRRALKY 33

RESULT 11
089898 ID 089898 PRELIMINARY; PRT; 54 AA.
AC 089898:
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Large T and small t antigens (fragment).
OS Buderiglar fledgling disease virus (BFDV).
OC Viruses; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.
CX NCBI_TaxId=10625;
RN [1]
RE SEQUENCE FROM N.A.
RP
RC STRAIN=BDSAB1-A;
RA Phalen D.N., Wilson V.G., Gaskin J.M., Derr J.N., Graham D.L.;
RT "Genetic Diversity in 20 Variants of the Avian Polyomavirus.";
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF054412; AAC33636.1; -
DR InterPro: IPR001623; DnaJ_N.
DR Pfam; PF00226; DnaJ_1.
DR SMART; SM00271; DnaJ_1.
DR PROSITE; PS50076; DNaJ_2; 1.
FT NON TER 54 54_
SQ SEQUENCE 54 AA; 6077 MW; 5AF094925DE8C997 CRC64;

Query Match 51.7%; Score 30; DB 12; Length 54;
Best Local Similarity 54.5%; Pred. No. 14e+02;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0

Oy      2 RLAIRRLALRY 12
       : | | | | :
Db      23 KTAAYRRALKY 33

RESULT 12
089899 ID 089899 PRELIMINARY; PRT; 54 AA.

```

[illegible][illegible]

OS Budgerigar fledgling disease virus (BFDV).
OC Viruses; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.
OX NCBI_TaxID=10625;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ECFL91;

RA Phalen D.N., Wilson V.G., Gaskin J.M., Derr J.N., Graham D.L.;

RT "Genetic Diversity in 20 Variants of the Avian Polyomavirus.;"

RL Submitted (MAR-1998) to the EMBL/GenBank/DDBJ databases.

DR EMBL; AF054415; AAC33639.1; -

DR InterPro; IPR001623; DnaJ_N.

DR Pfam; PF00226; DnaJ; 1.

DR SMART; SM00271; DnaJ; 1.

DR PROSITE; PS50076; DnaJ_2; 1.

FT NON_TER 54

SQ SEQUENCE 54 AA; 6077 MW; 5AF094925DE8C997 CRC64;

Query Match 51.7%; Score 30; DB 12; Length 54;

Best Local Similarity 54.5%; Pred. No. 1.4e+02;

Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 RLAIIRIALRY 12

Db 23 KTAVRTALKY 33

RESULT 15

O89902

ID O89902 PRELIMINARY; PRT; 54 AA.

AC O89902;

DT 01-NOV-1998 (TrEMBLrel. 08, Created)

DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)

DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)

DE Large t and small t antigens (Fragment).

OS Budgerigar fledgling disease virus (BFDV).

OC Viruses; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.

OX NCBI_TaxID=10625;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=LBFL93;

RA Phalen D.N., Wilson V.G., Gaskin J.M., Derr J.N., Graham D.L.;

RT "Genetic Diversity in 20 Variants of the Avian Polyomavirus.;"

RL Submitted (MAR-1998) to the EMBL/GenBank/DDBJ databases.

DR EMBL; AF054416; AAC33640.1; -

DR InterPro; IPR001623; DnaJ_N.

DR Pfam; PF00226; DnaJ; 1.

DR SMART; SM00271; DnaJ; 1.

DR PROSITE; PS50076; DnaJ_2; 1.

FT NON_TER 54

SQ SEQUENCE 54 AA; 6077 MW; 5AF094925DE8C997 CRC64;

Query Match 51.7%; Score 30; DB 12; Length 54;

Best Local Similarity 54.5%; Pred. No. 1.4e+02;

Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 RLAIIRIALRY 12

Db 23 KTAVRTALKY 33

Search completed: December 19, 2002, 16:21:50

Job time : 22.1765 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 19, 2002, 16:03:12 ; Search time 5.47059 Seconds
(without alignments)
90.980 Million cell updates/sec

Title: US-08-653-294C-36
Perfect score: 58
Sequence: 1 YRLAIRIALRY 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 5116

Minimum DB seq length: 0
Maximum DB seq length: 60

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	28	48.3	60	RS21_MYCPN	P57079 mycoplasma
2	27	46.6	23	TYPH_LACCA	P19663 lactobacill
3	27	46.6	42	REVS_BOVIN	P82708 bos taurus
4	27	46.6	47	H4Y_BLEJA	P80738 blepharisma
5	27	46.6	57	VSH_MUMPL	P28082 mumps virus
6	27	46.6	57	VSH_MUMPL	P28083 mumps virus
7	27	46.6	57	VSH_MUMPL	P22112 mumps virus
8	26	44.8	51	Y331_BORBU	O51309 borrelia bu
9	25	43.1	33	DEF1_MESAU	P81465 mesocricetu
10	25	43.1	37	PSBY_GUTH	O78433 guillardia
11	25	43.1	45	Y00D_BPT4	Q01434 bacterioph
12	25	43.1	57	VSH_MUMPT	P22113 mumps virus
13	24	41.4	49	Y42D_RHISN	P55731 rhizobium s
14	24	41.4	50	HSPI_PONPY	P35310 pongo pygma
15	24	41.4	57	VSH_MUMPA	P22111 mumps virus
16	24	41.4	57	VSH_MUMPA	P28087 mumps virus
17	23	39.7	36	PSBY_ODOSI	P49543 odontella s
18	23	39.7	37	RK36_PINTH	P41631 pinus thunb
19	23	39.7	47	RL34_BUCAI	P57129 buchnera ap
20	23	39.7	47	RL34_BUCAP	P29437 buchnera ap
21	23	39.7	47	RL40_METJA	P45058 methanococ
22	23	39.7	57	VSH_MUMPU	P20716 mumps virus
23	23	39.7	20	RECK_AZOVI	P37863 azotobacter
24	22	37.9	36	PMY_FETMA	P80024 petromyzon
25	22	37.9	37	RL36_CYACA	Q9tlu9 cyanidium c
26	22	37.9	45	SLIB_CYPCA	P42692 cyprinus ca
27	22	37.9	48	RL34_ODOSI	P49566 odontella s
28	22	37.9	49	HSPI_SAGIM	P24714 saginus im
29	22	37.9	50	HSPI_HYLLA	P24714 saginus im
30	22	37.9	51	HSPI_ALOSE	P35306 hylobates l
31	22	37.9	52	RL40_CAREL	P35302 alouatta se
32	22	37.9	52	RL40_HUMAN	P49632 caenorhabdi
33	22	37.9	56	NINE_BP82	P14793 homo sapien
					Q37871 bacterioph

34	22	37.9	56	1	NINE_ECOLI	Q47270 escherichia
35	22	37.9	56	1	RL33_TREPA	O83262 treponema p
36	22	37.9	58	1	HS33_HORSE	P15343 equus cabal
37	22	37.9	60	1	HSP1_CAEFU	P42131 caenolestes
38	21	36.2	24	1	HPTA_RABIT	P13571 oryctolagus
39	21	36.2	24	1	RS5_VIBPR	P52856 vibrio prot
40	21	36.2	30	1	PRT3_ONCMY	P02332 oncorhynch
41	21	36.2	30	1	PRT4_ONCMY	P02333 oncorhynch
42	21	36.2	31	1	DEF2_MESAU	P81466 mesocricetu
43	21	36.2	31	1	RECK_METCL	P37865 methylomona
44	21	36.2	33	1	DEF4_MESAU	P81468 mesocricetu
45	21	36.2	38	1	DCH5_MICSP	P00863 micrococcus

ALIGNMENTS

RESULT 1
RS21_MYCPN STANDARD; PRT; 60 AA.
AC P57079;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 30S ribosomal protein S21.
GN RPSU OR MPN296 OR MP539.1.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Hermann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae."
RT Nucleic Acids Res. 24:4420-4449 (1996).
RN [2]
RP IDENTIFICATION.
RX MEDLINE=20411492; PubMed=10954595;
RA Dandekar T., Huynen M., Regula J.T., Ueberle B., Zimmermann C.U.,
RA Andrade M.A., Doerks T., Sanchez-Pulido L., Snel B., Suyama M.,
RA Yuan Y.P., Herrmann R., Bork P.;
RT "Re-annotating the Mycoplasma pneumoniae genome sequence: adding
RT value, function and reading frames."
RL Nucleic Acids Res. 28:3278-3288 (2000).
CC -!- SIMILARITY: BELONGS TO THE S21P FAMILY OF RIBOSOMAL PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE000052; ; NOT ANNOTATED_CDS.
DR EMBL; AE000053; AAG34754.1; -
DR InterPro; IPR001911; Ribosomal S21.
DR PROSITE; PS01181; RIBOSOMAL S21; FALSE_NEG.
KW Ribosomal protein; Complete proteome.
SQ SEQUENCE 60 AA; 7515 MW; 30D87DCF4635F003 CRC64;

Query Match 48.3%; Score 28; DB 1; Length 60;
Best Local Similarity 60.0%; Pred. No. 40;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 LAIRIALRY 12
DB 23 LEIRLAQRH 32

```

RESULT 2
TYPH_LACCA STANDARD; PRT; 23 AA.
ID TYPH_LACCA
AC P19663;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Thymidine phosphorylase (EC 2.4.2.4) (TDRPASE) (Fragment).
GN DEOA.
OS Lactobacillus casei.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=1562;
RN [1]
RP SEQUENCE.
RC STRAIN=ATCC 7469;
RX MEDLINE=90381286; PubMed=2119230;
RA Avraham Y., Grossowicz N., Yashphe J.;
RT "Purification and characterization of uridine and thymidine
   phosphorylase from Lactobacillus casei.";
RL Biochim. Biophys. Acta 1040:287-293(1990).
CC -1- FUNCTION: THE ENZYMES WHICH CATALYZE THE REVERSIBLE PHOSPHORYLOSIS
   OF PYRIMIDINE NUCLEOSIDES ARE INVOLVED IN THE DEGRADATION OF THESE
   COMPOUNDS AND IN THEIR UTILIZATION AS CARBON AND ENERGY SOURCES,
   OR IN THE RESCUE OF PYRIMIDINE BASES FOR NUCLEOTIDE SYNTHESIS.
CC -1- CATALYTIC ACTIVITY: Thymidine + phosphate = thymine + 2-deoxy-D-
   ribose 1-phosphate.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SIMILARITY: BELONGS TO THE THYMIDINE/PYRIMIDINE-NUCLEOSIDE
   PHOSPHORYLASES FAMILY.
CC PIR: S11384; S11384.
DR InterPro: IPR000173; GAP_dhdrogenase.
DR InterPro: IPR000053; Thymid_phosphils.
DR Pfam: PF00044; spdh; 1.
DR PROSITE: PS00647; THYMID_PHOSPHORYLASE; PARTIAL.
KW Transferase; Glycosyltransferase.
FT NON_TER 23
SQ SEQUENCE 23 AA; 2752 MW; 8705C9C4D82C1AD7 CRC64;

Query Match 46.6%; Score 27; DB 1; Length 23;
Best Local Similarity 85.7%; Pred. No. 22;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 RLAIIRRI 8
Db 14 RLAIIRRI 20

RESULT 3
RETS_BOVIN STANDARD; PRT; 42 AA.
ID RETS_BOVIN
AC P82708;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Retinol-binding protein III, cellular (CRBP-III) (Fragment).
GN RBP5.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE, AND FUNCTION.
RC TISSUE=Kidney;
RX MEDLINE=21173623; PubMed=11274389;
RA Polli C., Calderone V., Otonello S., Bolchi A., Zanotti G.,
RT "Identification, retinoid binding and X-ray analysis of a human
   retinol-binding protein.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:3710-3715(2001).
CC -1- FUNCTION: Intracellular transport of retinol.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).

```

```

CC -1- TISSUE SPECIFICITY: KIDNEY.
CC -1- SIMILARITY: BELONGS TO THE FABP/P2/CRBP/CRABP FAMILY OF
   TRANSPORTERS.
DR HSSP; P82980; 1GGU.
DR InterPro: IPR000463; Fatty_acid_BP.
DR InterPro: IPR000566; Lipocln_cytfabp.
DR Pfam: PF00061; lipocalin; 1.
DR PROSITE: PS00214; FABP; FALSE_NEG.
KM Vitamin A; Retinol-binding; Transport.
FT NON_TER 42
SQ SEQUENCE 42 AA; 4892 MW; ACB4F139FPDD7F09 CRC64;

Query Match 46.6%; Score 27; DB 1; Length 42;
Best Local Similarity 62.5%; Pred. No. 43;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 LAIRRIAL 10
Db 27 MALRIKIAL 34

RESULT 4
HAY_BLEJA STANDARD; PRT; 47 AA.
ID HAY_BLEJA
AC P80738;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Histone H4 (Fragment).
OS Blepharisma japonicum.
OC Eukaryota; Alveolata; Ciliophora; Heterotricha; Heterotrichida;
OC Blepharisma
OX NCBI_TaxID=5961;
RN [1]
RP SEQUENCE.
RC STRAIN=A5-3;
RX MEDLINE=97257506; PubMed=9103980;
RA Salvini M., Bini E., Santucci A., Batistoni R.;
RT "H4 histone in the macronucleus of Blepharisma japonicum (Protozoa,
   Ciliophora, Heterotrichida)."
RL FEWS Microbiol. Lett. 149:93-98(1997).
CC -1- FUNCTION: HISTONE H4, ALONG WITH HISTONE H3, PLAYS A CENTRAL ROLE
   IN NUCLEOSOME FORMATION.
CC -1- SUBUNIT: THE NUCLEOSOME IS AN OCTAMER CONTAINING TWO MOLECULES OF
   H2A, H2B, H3, AND H4, WHICH WRAP APPROXIMATIVELY 146 BP OF DNA.
CC -1- SIMILARITY: BELONGS TO THE HISTONE H4 FAMILY.
DR InterPro: IPR001951; Histone_H4.
DR ProDom: PD001827; Histone_H4; 1.
DR PROSITE: PS00047; HISTONE_H4; PARTIAL.
KM Chromosomal protein; Nucleosome core; Nuclear protein; DNA-binding.
FT NON_TER 47
SQ SEQUENCE 47 AA; 4961 MW; D743BEA73B76BADA CRC64;

Query Match 46.6%; Score 27; DB 1; Length 47;
Best Local Similarity 75.0%; Pred. No. 49;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 AIRRIALR 11
Db 36 AIRRIARR 43

RESULT 5
VSH_MUMPB STANDARD; PRT; 57 AA.
ID VSH_MUMPB
AC P28082;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Small hydrophobic protein.
GN SH.
OS Mumps virus (strain Belfast).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;

```

```
OC Paramyxoviridae; Paramyxovirinae; Rubulavirus.
OX NCBI_TaxID=11166;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93168036; PubMed=8435047;
RA Yeo R.P., Afzal M.A., Forsey T., Rima B.K.;
RT "Identification of a new mumps virus lineage by nucleotide sequence
RT analysis of the SH gene of ten different strains.";
RL Arch. Virol. 128:371-377(1993).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X63709; CAM45242.1; --
DR FIR; S19866; SHNZBF.
DR InterPro; IPR001477; SH.
DR Pfam; PF01445; SH; 1.
DR ProDom; PD001504; SH; 1.
DR Transmembrane.
KW Transmembrane.
FT TRANSMEM 8 32
SQ SEQUENCE 57 AA; 6814 MW; 8BF2F917B4A8EF3C CRC64;
-----
Query Match 46.6%; Score 27; DB 1; Length 57;
Best Local Similarity 50.0%; Pred. No. 61;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 1 YRLAIRRIAL 10
DB 34 YKTAVERHAAL 43
-----
RESULT 6
VSH MUMPL STANDARD; PRT; 57 AA.
AC P28083;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Small hydrophobic protein.
GN SH.
OS Mumps virus (strain Bristol 1).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Rubulavirus.
OX NCBI_TaxID=11170;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93168036; PubMed=8435047;
RA Yeo R.P., Afzal M.A., Forsey T., Rima B.K.;
RT "Identification of a new mumps virus lineage by nucleotide sequence
RT analysis of the SH gene of ten different strains.";
RL Arch. Virol. 128:371-377(1993).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X63713; CAM45246.1; --
DR FIR; S19870; SHNZBL.
DR InterPro; IPR001477; SH.
DR Pfam; PF01445; SH; 1.
DR ProDom; PD001504; SH; 1.
DR Transmembrane.
KW Transmembrane.
FT TRANSMEM 8 32
SQ SEQUENCE 57 AA; 6753 MW; 8EF2EC57BA4FBC1D CRC64;
-----
OC Paramyxoviridae; Paramyxovirinae; Rubulavirus.
OX NCBI_TaxID=11166;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93168036; PubMed=8435047;
RA Yeo R.P., Afzal M.A., Forsey T., Rima B.K.;
RT "Identification of a new mumps virus lineage by nucleotide sequence
RT analysis of the SH gene of ten different strains.";
RL Arch. Virol. 128:371-377(1993).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X63709; CAM45242.1; --
DR FIR; S19866; SHNZBF.
DR InterPro; IPR001477; SH.
DR Pfam; PF01445; SH; 1.
DR ProDom; PD001504; SH; 1.
DR Transmembrane.
KW Transmembrane.
FT TRANSMEM 8 32
SQ SEQUENCE 57 AA; 6814 MW; 8BF2F917B4A8EF3C CRC64;
-----
Query Match 46.6%; Score 27; DB 1; Length 57;
Best Local Similarity 50.0%; Pred. No. 61;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 1 YRLAIRRIAL 10
DB 34 YKTAVERHAAL 43
-----
RESULT 7
VSH MUMPM STANDARD; PRT; 57 AA.
AC P22112;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 01-AUG-1991 (Rel. 19, Last annotation update)
DE Small hydrophobic protein.
GN SH.
OS Mumps virus (strain Miyahara vaccine).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Rubulavirus.
OX NCBI_TaxID=11171;
RN [1]
RP SEQUENCE FROM N.A.
RA Takeuchi K.;
RL Submitted (NOV-1990) to the EMBL/GenBank/DDBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D90234; BAAL4282.1; --
DR FIR; JU0304; SHNZMS.
DR InterPro; IPR001477; SH.
DR Pfam; PF01445; SH; 1.
DR ProDom; PD001504; SH; 1.
DR Transmembrane.
KW Transmembrane.
FT TRANSMEM 8 32
SQ SEQUENCE 57 AA; 6621 MW; BC794E57A4A0F9B6 CRC64;
-----
Query Match 46.6%; Score 27; DB 1; Length 57;
Best Local Similarity 50.0%; Pred. No. 61;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 1 YRLAIRRIAL 10
DB 34 YKTAVERHAAL 43
-----
RESULT 8
Y331 BORBU STANDARD; PRT; 51 AA.
AC O51309;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein BB0331.
GN BB0331.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 35210 / B31;
RX MEDLINE=98065943; PubMed=9403685;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
```

RA Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
 RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
 RA Usterback T., Wathney L., McDonald L., Artlich P., Bowman C.,
 RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
 RA Smith H.O., Venter J.C.;
 RT "Genomic sequence of a Lyme disease spirochaete, *Borrelia burgdorferi*,"
 RL Nature 390:580-586 (1997).

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 CC EMBL; AF001140; AAC66728.1; -.
 DR TIGR; BB0331; -.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 51 AA; 6444 MW; ACECTE55B5A8E0DA CRC64;

 Query Match 44.8%; Score 26; DB 1; Length 51;
 Best Local Similarity 60.0%; Pred. No. 86;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

 QY 1 YRLAIRRL 10
 Db 21 YSLDRLRIFL 30

 RESULT 9
 DEFL_MESAU STANDARD; PRT; 33 AA.
 ID P81465;
 AC 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Neutrophil defensin 1 (HAMP-1).
 OS Mesocricetus auratus (Golden hamster).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
 CC Mesocricetus.
 CC NCBI_TaxID=10036;
 CC [1]
 RP SEQUENCE.
 RX MEDLINE=97045125; PubMed=8890190;
 RA Mak P., Wojcik K., Thogersen I.B., Dublin A.;
 RT "Isolation, antimicrobial activities, and primary structures of
 RT hamster neutrophil defensins,"
 RL Infect. Immun. 64:4444-4449 (1996).
 CC -1- FUNCTION: ANTI-FUNGI AND BACTERICIDAL ACTIVITY, GREATER AGAINST
 CC GRAM-POSITIVE BACTERIA.
 CC -1- SIMILARITY: BELONGS TO THE CORTICOSTATIN/DEFENSIN FAMILY.
 DR InterPro; IPR001271; Defensin_mammal.
 DR Pfam; PF00323; defensins; 1.
 DR SMART; SM00048; DEFENSIN; 1.
 DR PROSITE; PS00269; DEFENSIN; 1.
 KW Defensin; Antibiotic; Fungicide.
 FT DISULFID 3 31 BY SIMILARITY.
 FT DISULFID 5 20 BY SIMILARITY.
 FT DISULFID 10 30 BY SIMILARITY.
 SQ SEQUENCE 33 AA; 4014 MW; 07DBDFDCFC1BDA CRC64;

 Query Match 43.1%; Score 25; DB 1; Length 33;
 Best Local Similarity 71.4%; Pred. No. 84;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

 QY 1 YRLAIR 7
 Db 27 YRLCGR 33

RESULT 10
 ID PSBY_GUTH STANDARD; PRT; 37 AA.
 AC 078433;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Photosystem II protein Y.
 GS PSBY OR YCF32.
 OS Guillardia theta (Cryptomonas ph.).
 CC Chloroplast.
 CC Eukaryota; Cryptophyta; Cryptomonadaceae; Guillardia.
 CC NCBI_TaxID=55529;
 CC [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99128221; PubMed=9929392;
 RA Douglas S.E., Penny S.L.;
 RT "The plastid genome of the cryptophyte alga, *Guillardia theta*:
 RT complete sequence and conserved synteny groups confirm its common
 RT ancestry with red algae,"
 RL J. Mol. Evol. 48:236-244 (1999).
 CC -1- FUNCTION: MANGANESE-BINDING POLYPEPTIDE WITH L-ARGININE
 CC METABOLIZING ENZYME ACTIVITY. COMPONENT OF THE CORE OF PHOTOSYSTEM
 CC II (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Chloroplast thylakoid membrane.
 CC -1- SIMILARITY: BELONGS TO THE PSBY FAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 CC EMBL; AF041468; AAC35618.1; -.
 DR Photosystem II; Chloroplast; Transmembrane; Thylakoid.
 KM DOMAIN 1 6 LUMENAL (POTENTIAL).
 FT TRANSMEM 7 23 POTENTIAL.
 FT DOMAIN 24 37 STROMAL (POTENTIAL).
 SQ SEQUENCE 37 AA; 4094 MW; 55A1B1ED934A655D CRC64;

 Query Match 43.1%; Score 25; DB 1; Length 37;
 Best Local Similarity 85.7%; Pred. No. 96;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

 QY 5 IRLAIR 11
 Db 23 IRLAIR 29

 RESULT 11
 ID YOOD_BPT4 STANDARD; PRT; 45 AA.
 AC Q01434;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical 5.1 kDa protein in Gp39-comCA intergenic region.
 GN YOOD OR 39.2 OR COMCA.-1.
 OS Bacteriophage T4.
 CC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
 CC T4-like viruses.
 CC NCBI_TaxID=10665;
 CC [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93015705; PubMed=1400206;
 RA Sanson B., Uzan M.;
 RT "Sequence and characterization of the bacteriophage T4 comC alpha
 RT gene product, a possible transcription antitermination factor,"
 RL J. Bacteriol. 174:6539-6547 (1992).
 CC [2]
 RP SEQUENCE FROM N.A.

RA Kutter E., Ariesaka F., Kunisawa T., Teugita A., Mosig G.,
 RA Mesyanzhinov V., Ruger W., Stidham T., Thomas E.,
 RT "Bacteriophage T4 genome analysis."
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR ENBL; M89919; AAA32486.1; --
 DR ENBL; AF158101; AAD42464.1; --
 DR PIR; D45731; D45731.
 KW Hypothetical protein.
 SQ SEQUENCE 45 AA; 5107 MW; 4E550618271A1257 CRC64;

Query Match 43.1%; Score 25; DB 1; Length 45;
 Best Local Similarity 44.4%; Pred. No. 1.2e+02;
 Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 YRLAIRRIA 9
 Db 35 HRLAVRLVS 43

RESULT 12
 VSH MUMPT
 ID VSH_MUMPT STANDARD; PRT; 57 AA.
 AC F22113;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 01-AUG-1991 (Rel. 19, Last annotation update)
 DE Small hydrophobic protein.
 GN SH.
 OS Mumps virus (strain Takahashi).
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
 OC Paramyxoviridae; Paramyxovirinae; Rubulaviruses.
 OX NCBI_TaxID=11174;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Takeuchi K.;
 RL Submitted (NOV-1990) to the EMBL/GenBank/DBJ databases.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR ENBL; D90235; BAA14283.1; --
 DR PIR; JU0309; SHNZMT.
 DR InterPro; IPR001477; SH.
 DR Pfam; PF01445; SH; 1.
 DR ProDom; PD001504; SH; 1.
 KW Transmembrane.
 FT TRANSMEM 8 32 POTENTIAL.
 SQ SEQUENCE 57 AA; 6745 MW; EC965B46AB45ECF0 CRC64;

Query Match 43.1%; Score 25; DB 1; Length 57;
 Best Local Similarity 50.0%; Pred. No. 1.6e+02;
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 YRLAIRRIA 10
 Db 34 YKTAMRHAAL 43

RESULT 13

Y4ZD RHISN
 ID Y4ZD_RHISN STANDARD; PRT; 49 AA.
 AC P55731;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Hypothetical 5.5 kDa protein Y4ZD.
 GN Y4ZD.
 OS Rhizobium sp. (strain NGR234).
 OG Plasmid sym pNGR234.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Rhizobium.
 OX NCBI_TaxID=394;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97305956; PubMed=9163424;
 RA Freiberg C.A., Fellay R., Bairoch A., Broughton W.J., Rosenthal A.,
 RA Perret X.;
 RT "Molecular basis of symbiosis between Rhizobium and legumes.";
 RL Nature 387:394-401(1997).
 CC -!- SIMILARITY: NONE OBVIOUS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR ENBL; AE000109; AAB91962.1; --
 KW Hypothetical protein; Plasmid.
 SQ SEQUENCE 49 AA; 5514 MW; 6B1CF7141547BD0D CRC64;

Query Match 41.4%; Score 24; DB 1; Length 49;
 Best Local Similarity 71.4%; Pred. No. 2.1e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 IRRIALR 11
 Db 5 IQRISLR 11

RESULT 14
 HSP1_PONPY
 ID HSP1_PONPY STANDARD; PRT; 50 AA.
 AC P35110;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE SpERM protamine P1 (Cysteine-rich protamine).
 GN PRM1.
 OS Pongo pygmaeus (Orangutan).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
 OX NCBI_TaxID=9600;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94141943; PubMed=8308910;
 RA Rettief J.D., Winkfein R.J., Dixon G.H., Adroer R., Queralt R.,
 RA Ballabriga J., Oliva R.;
 RT "Evolution of protamine P1 genes in primates.";
 RL J. Mol. Evol. 37:426-434(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20123244; PubMed=10659848;
 RA Wyckoff G.J., Wang W., Wu C.-I.;
 RT "Rapid evolution of male reproductive genes in the descent of man.";
 RL Nature 403:304-309(2000).
 CC -!- FUNCTION: PROTAMINES SUBSTITUTE FOR HISTONES IN THE CHROMATIN OF
 CC SPERM DURING THE HAPLOID PHASE OF SPERMATOGENESIS. THEY COMPACT
 CC SPERM DNA INTO A HIGHLY CONDENSED, STABLE AND INACTIVE COMPLEX.
 CC -!- SUBUNIT: CROSS-LINKED BY INTERCHAIN DISULFIDE BONDS AROUND THE

CC DNA-HELIX (BY SIMILARITY).

CC - SUBCELLULAR LOCATION: Nuclear.

CC - TISSUE SPECIFICITY: TESTIS.

CC - SIMILARITY: BELONGS TO THE PROTAMINE P1 FAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

CC EMBL; L14589; AAA36946.1; -

DR EMBL; AF215710; AAF34623.1; -

DR InterPro; IPR000221; Protamine_P1.

DR Pfam; PF00260; protamine_P1; 1

DR PROSITE; PS00048; PROTAMINE_P1; 1

KM Chromosomal protein; Nucleosome core; Spermatogenesis; DNA-binding;

KW Testis; DNA condensation; Nuclear protein.

FT INIT MET 0 0 BY SIMILARITY.

SO SEQUENCE 50 AA; 6660 MW; D9935BFA11B5BF78 CRC64;

Query Match 41.4%; Score 24; DB 1; Length 50;

Best Local Similarity 71.4%; Pred. No. 2.1e+02; Indels 0; Gaps 0;

Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 YRLAIR 7

Db 43 YRLRCR 49

RESULT 15

ID VSH_MUMPA STANDARD; PRT; 57 AA.

AC P22111;

DT 01-AUG-1991 (Rel. 19, Created)

DT 01-AUG-1991 (Rel. 19, Last sequence update)

DT 01-AUG-1991 (Rel. 19, Last annotation update)

DE Small hydrophobic protein.

GN SH.

OS Mumps virus (strain Matsuyama).

OC Viruses; ssRNA negative-strand viruses; Mononegavirales;

OC Paramyxoviridae; Paramyxovirinae; Rubulavirus.

OX NCBI_TaxID=11165;

RM [1]

RP SEQUENCE FROM N.A.

RA Takeuchi K.;

RL Submitted (NOV-1990) to the EMBL/GenBank/DBJ databases.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

CC EMBL; D90233; BAA14281.1; -

DR PIR; J00307; SHNZMM.

DR InterPro; IPR001477; SH.

DR Pfam; PF01445; SH; 1.

DR ProDom; PD001504; SH; 1.

KW Transmembrane.

FT TRANSMEM 8 32 POTENTIAL.

SO SEQUENCE 57 AA; 6697 MW; 1D94B57AAS5ECF0 CRC64;

Query Match 41.4%; Score 24; DB 1; Length 57;

Best Local Similarity 40.0%; Pred. No. 2.5e+02; Indels 0; Gaps 0;

Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 YRLAIRRL 10

Db 43 YRLRCR 49

Db 34 YKTAVRHASL 43

Search completed: December 19, 2002, 16:19:42
 Job time : 5.47059 secs

GenCore version 5.1.1.3

Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 19, 2002, 16:16:02 ; Search time 10.4118 Seconds
(without alignments)
110:799 Million cell updates/sec

Title: US-08-653-294C-36

Perfect score: 58

Sequence: 1 YRLAIRRLRY 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 15911

Minimum DB seq length: 0

Maximum DB seq length: 60

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 73:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30	51.7	53	2 C82714	50S ribosomal prot
2	28	48.3	52	2 D84640	hypothetical prote
3	28	48.3	58	2 A86702	hypothetical prote
4	27	46.6	24	2 S11384	thymidine phosphor
5	27	46.6	34	2 A60475	glyceraldhyde-3-p
6	27	46.6	57	1 SHNZMS	small hydrophobic
7	27	46.6	57	1 SHNZBF	small hydrophobic
8	26	44.8	51	2 B70141	hypothetical prote
9	26	44.8	52	2 AF2002	hypothetical prote
10	26	44.8	55	2 E82080	hypothetical prote
11	25	43.1	28	2 PH1335	Ig heavy chain DJ
12	25	43.1	33	2 A82288	hypothetical prote
13	25	43.1	38	2 AB0747	hypothetical prote
14	25	43.1	45	2 D45731	gene 39.2 protein
15	25	43.1	52	2 T31016	hypothetical prote
16	25	43.1	57	1 SHNZMT	small hydrophobic
17	25	43.1	57	2 JQ2368	small hydrophobic
18	25	43.1	58	2 AB1589	hypothetical prote
19	25	43.1	58	2 A11650	hypothetical prote
20	25	43.1	59	2 T12018	plex protein - Pse
21	25	43.1	59	2 AE1889	hypothetical prote
22	24	41.4	39	2 C55995	prostaglandin E2 r
23	24	41.4	39	2 G90716	probable RNA [impo
24	24	41.4	39	2 G64801	hypothetical prote
25	24	41.4	44	2 S29785	NADH oxidase (H2O2
26	24	41.4	48	2 D55995	prostaglandin E2 r
27	24	41.4	52	2 C84186	hypothetical prote
28	24	41.4	53	2 G82813	hypothetical prote
29	24	41.4	54	2 D82574	hypothetical prote

30	24	41.4	57	1 SHNZMM	small hydrophobic
31	24	41.4	57	2 C41476	ribonucleoside-dip
32	24	41.4	57	2 S24827	small hydrophobic
33	23	39.7	34	2 JC4218	phosphofructokinase
34	23	39.7	36	2 S78239	ycf32 protein - Od
35	23	39.7	37	2 T07519	ribosomal protein
36	23	39.7	41	2 T07461	hypothetical prote
37	23	39.7	44	2 A81550	hypothetical prote
38	23	39.7	47	1 JG1154	ribosomal protein
39	23	39.7	47	2 E84931	50S ribosomal prot
40	23	39.7	47	2 C64388	ribosomal protein
41	23	39.7	49	2 S58608	hypothetical prote
42	23	39.7	52	2 F70209	hypothetical prote
43	23	39.7	53	2 B69161	hypothetical prote
44	23	39.7	55	2 G98149	hypothetical prote
45	23	39.7	57	1 SHNZMU	small hydrophobic

ALIGNMENTS

RESULT 1

C82714

50S ribosomal protein L30 XF1170 [imported] - Xylella fastidiosa (strain 9a5c)

C:Species: Xylella fastidiosa

C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 17-Nov-2000

C:Accession: C82714

R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen

Nature 406, 151-157, 2000

A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.

A:Reference number: A82515; MUID:20365717; PMID:10910347

A:Note: for a complete list of authors see reference number A59328 below

A:Accession: C82714

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-53 <SIM>

A:Cross-references: GB:AB003951; GB:AE003849; NID:99106131; PIDN:AAF83980.1; GSPDB:GN001

R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A

Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Carraro, L.E.A.; Carraro, D.M.; Carrer, H

as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.

submitted to GenBank, June 2000

A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm

J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig

Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, P

, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A

Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawa

A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir

M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; V

A:Reference number: A59328

A:Contents: annotation

C:Genetics:

A:Gene: XF1170

C:Superfamily: Escherichia coli ribosomal protein L30

Query Match

Best Local Similarity 51.7%; Score 30; DB 2; Length 53;

Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 YRLAIRRLRY 11

:|:|:|:|:|

Db 14 HRLSVRLGLR 24

RESULT 2

D84640

hypothetical protein At2g24780 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001

C:Accession: D84640

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.

euss, D.; Niernan, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
 Nature 402, 761-768, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.
 A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: D84640
 A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-52 <STO>

A:Cross-references: GB:AE002093; NID:g4559373; PIDN:AAD23034.1; GSPDB:GN00139

C:Genetics:

A:Gene: At2g24780

A:Map position: 2

Query Match 48.3%; Score 28; DB 2; Length 52;
 Best Local Similarity 60.0%; Pred. No. 1.5e+02;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 RLAIRRIAR 11
 Db 19 KLIRIRIR 28

RESULT 3

A86702
 hypothetical protein ygaF [imported] - *Lactococcus lactis* subsp. *lactis* (strain IIL403)

C:Species: *Lactococcus lactis* subsp. *lactis*

C>Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001

C:Accession: A86702

R:Botolin, A.; Wincker, P.; Mauger, S.; Jallion, O.; Malarme, K.; Weissenbach, J.; Ehrlich

Genome Res. 11, 731-753, 2001

A:Title: The complete genome sequence of the lactic acid bacterium *Lactococcus lactis* s.s.

A:Reference number: A86625; MUID:21235186; PMID:11337471

A:Accession: A86702

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-58 <STO>

A:Cross-references: GB:AE005176; PID:g12723514; PIDN:AAK04715.1; GSPDB:GN00146

A:Experimental source: strain IIL403

C:Genetics:

A:Gene: ygaF

Query Match 48.3%; Score 28; DB 2; Length 58;
 Best Local Similarity 60.0%; Pred. No. 1.7e+02;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 3 LAIRRIARY 12
 Db 38 LVRRIPLAY 47

RESULT 4

S11384
 chymidine phosphorylase (EC 2.4.2.4) - *Lactobacillus casei* (fragment)

N:Alternate names: pyrimidine phosphorylase

C:Species: *Lactobacillus casei*

C>Date: 19-Mar-1997 #sequence_revision 26-Feb-1998 #text_change 21-Nov-1998

C:Accession: S11384

R:Avraham, Y.; Grossowicz, N.; Yashphe, J.

Biochim. Biophys. Acta 1040, 287-293, 1990

A:Title: Purification and characterization of uridine and chymidine phosphorylase from *L.*

A:Reference number: S11383; MUID:90381286; PMID:2119230

A:Accession: S11384

A:Molecule type: protein

A:Residues: 1-24 <AVR>

A:Note: this protein was shown to possess chymidine phosphorylase activity and not glycyl

C:Superfamily: glyceraldehyde-3-phosphate dehydrogenase

C:Keywords: glycosyltransferase; homodimer; pentosyltransferase

Query Match 46.6%; Score 27; DB 2; Length 24;
 Best Local Similarity 85.7%; Pred. No. 1.1e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 RLAIRRI 8
 Db 15 RLAFRRI 21

RESULT 5

A60475
 glyceraldehyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12), pentalenolacton

C:Species: *Streptomyces arenae*

C>Date: 20-Feb-1993 #sequence_revision 20-Feb-1993 #text_change 03-Jun-2002

C:Accession: A60475; A44754

R:Freihich, K.U.; Wiedmann, M.; Lottspeich, F.; Mecke, D.

J. Bacteriol. 171, 6596-6702, 1989

A:Title: Substitution of a pentalenolactone-sensitive glyceraldehyde-3-phosphate dehydro

ae.

A:Reference number: A60475; MUID:90078116; PMID:2592349

A:Accession: A60475

A:Molecule type: protein

A:Residues: 1-34 <PRO>

C:Superfamily: glyceraldehyde-3-phosphate dehydrogenase

C:Keywords: oxidoreductase

Query Match 46.6%; Score 27; DB 2; Length 34;
 Best Local Similarity 85.7%; Pred. No. 1.6e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 RLAIRRI 8
 Db 14 RLAFRRI 20

RESULT 6

SHNZMS
 small hydrophobic protein - mumps virus (strain Miyahara)

C:Species: mumps virus

C>Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 16-Jun-2000

C:Accession: J00304

R:Takeuchi, K.

submitted to JIPID, November 1990

A:Reference number: J00304

A:Accession: J00304

A:Molecule type: mRNA

A:Residues: 1-57 <TAK>

A:Cross-references: GB:D90234; NID:g222162; PIDN:BAI4282.1; PID:g222163

C:Superfamily: mumps virus small hydrophobic protein

C:Keywords: transmembrane protein

F.13-29/Domain: transmembrane #status predicted <TM1>

Query Match 46.6%; Score 27; DB 1; Length 57;
 Best Local Similarity 50.0%; Pred. No. 2.6e+02;
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 YRLAIRRIAL 10
 Db 34 YXTAVRHALL 43

RESULT 7

SHNZBF
 small hydrophobic protein - mumps virus (strain Belfaer)

C:Species: mumps virus

C>Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 16-Jul-1999

C:Accession: S19866

R:Yeo, R.P.; Afzal, M.A.; Forsey, T.; Rina, B.K.

submitted to the EMBL Data Library, January 1992

A:Description: Nucleotide sequence analysis of the SH gene of mumps virus reveals severa

A:Reference number: S19866

A:Accession: S19866

A:Molecule type: DNA

A:Residues: 1-57 <YEO>

A:Cross-references: EMBL:X63709; NID:g60596; PIDN:CAA45242.1; PID:g60597

C:Genetics:

A:Gene: SH

C:Superfamily: mumps virus small hydrophobic protein
 C:Keywords: transmembrane protein
 F:8-29/Domain: transmembrane #status predicted <TMM>

Query Match 46.6%; Score 27; DB 1; Length 57;
 Best Local Similarity 50.0%; Pred. No. 2.6e+02;
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 YRLAIRRIAL 10
 DB 34 YKTAVRHAAL 43

RESULT 8

B70141

hypothetical protein BB0331 - Lyme disease spirochete

C:Species: Borrelia burgdorferi (Lyme disease spirochete)

C>Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 08-Oct-1999

C:Accession: B70141

R:Fraser, C.M.; Carjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White
 son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,
 ; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
 Nature 390, 580-586, 1997

A:Authors: Smith, H.O.; Venter, J.C.

A>Title: Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.

A:Reference number: A70100; MUID:98065943; PMID:9403685

A:Accession: B70141

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-51 <KLE>

A:Cross-references: GB:AE001140; GB:AE000783; NID:g2688233; PIDN:AAC66728.1; PID:g268824

A:Experimental source: strain B31

Query Match

44.8%; Score 26; DB 2; Length 51;

Best Local Similarity 60.0%; Pred. No. 3.6e+02;

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 YRLAIRRIAL 10

DB 21 YSLDLRRIFL 30

RESULT 9

A2002

hypothetical protein asr1572 [imported] - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp.

A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002

C:Accession: AF2002

R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
 Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
 DNA Res. 8, 205-213, 2001

A>Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana

A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: AF2002

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-52 <KUR>

A:Cross-references: GB:BA000019; PIDN:BA877938.1; PID:g17135392; GSPDB:GN00179

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: asr1572

Query Match

44.8%; Score 26; DB 2; Length 52;

Best Local Similarity 50.0%; Pred. No. 3.7e+02;

Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 LAIRRIALRY 12

DB 13 LKLRRAILKY 22

RESULT 10

E82080

hypothetical protein VC2393 [imported] - Vibrio cholerae (strain N16961 serogroup O1)

C:Species: Vibrio cholerae

C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001

C:Accession: E82080

R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
 chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, I.;
 l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A>Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A:Reference number: A82035; MUID:20406833; PMID:10952301

A:Accession: E82080

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-55 <HEI>

A:Cross-references: GB:AE004310; GB:AE003852; NID:g9655167; PIDN:AAF95536.1; GSPDB:GN001

A:Experimental source: serogroup O1; strain N16961; biotype El Tor

C:Genetics:

A:Gene: VC2393

A:Map position: 1

Query Match

44.8%; Score 26; DB 2; Length 55;

Best Local Similarity 62.5%; Pred. No. 3.9e+02;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 LAIRRIAL 10

DB 37 LGIKRVAL 44

RESULT 11

PH1335

ig heavy chain DJ region (clone C155-137) - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999

C:Accession: PH1335

R:Wasserman, R.; Galili, N.; Ito, Y.; Reichard, B.A.; Shane, S.; Rovera, G.

J. Exp. Med. 176, 1577-1581, 1992

A>Title: Predominance of fetal type DJH joining in young children with B precursor lymph

A:Reference number: PH1302; MUID:93094761; PMID:1460419

A:Accession: PH1335

A:Molecule type: DNA

A:Residues: 1-28 <WAS>

A:Note: the authors translated the stop codon for residue 4 as X

C:Keywords: heterotetramer; immunoglobulin

Query Match

43.1%; Score 25; DB 2; Length 28;

Best Local Similarity 50.0%; Pred. No. 3.2e+02;

Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 LAIRRIALRY 12

DB 11 VIRULLLRY 20

RESULT 12

A82288

hypothetical protein VC0735 [imported] - Vibrio cholerae (strain N16961 serogroup O1)

C:Species: Vibrio cholerae

C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001

C:Accession: A82288

R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
 chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, I.;
 l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A>Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A:Reference number: A82035; MUID:20406833; PMID:10952301

A:Accession: A82288

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-33 <HEI>

A:Cross-references: GB:AE004159; GB:AE003852; NID:g9655167; PIDN:AAF93900.1; GSPDB:GN001

A:Experimental source: serogroup O1; strain N16961; biotype El Tor

C:Genetics:
A:Gene: VC0735
A:Map position: 1

Query Match 43.1%; Score 25; DB 2; Length 33;
Best Local Similarity 50.0%; Pred. No. 3.7e+02;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 LAIRRIALRY 12
|||:|:
Db 15 LADRKVSRY 24

RESULT 13

AB0747
hypothetical protein STY2135 [imported] - *Salmonella enterica* subsp. *enterica* serovar Typhimurium
C:Species: *Salmonella enterica* subsp. *enterica* serovar Typhimurium
A:Note: this species has also been called *Salmonella typhimurium*
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 09-Nov-2001
C:Accession: AB0747
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, C.; Connor, T.; Connor, P.; Cronin, A.; Davies, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Park, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; A:Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serovar Typhimurium
A:Reference number: AB0502; PMID:11677608
A:Accession: AB0747
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-38 <PAR>
A:Cross-references: GB:AL513382; PDB:1G503172; GSPDB:GN00176
C:Genetics:
A:Gene: STY2135

Query Match 43.1%; Score 25; DB 2; Length 38;
Best Local Similarity 45.5%; Pred. No. 4.3e+02;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 RLAIIRIALRY 12
|||:|:
Db 17 RLAIIRIALRY 27

RESULT 14

D45731
gene 39.2 protein - phage T4
N:Alternate names: gp 39.2

C:Species: phage T4
A:Note: host *Escherichia coli*
C:Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 08-Oct-1999
C:Accession: J05054; D45731
R:Huang, W.M.

Nucleic Acids Res. 14, 7751-7765, 1986
A:Title: Nucleotide sequence of a type II DNA topoisomerase gene. Bacteriophage T4 gene 39.2

A:Reference number: A25763; MUID:87040739; PMID:3022233
A:Accession: J05054

A:Molecule type: DNA
A:Residues: 1-37 <HUA>

A:Cross-references: GB:X06220; NID:9728616
A:Note: this ORF is not annotated in GenBank entry MYT4G39

R:Sanison, B.; Uzan, M.

J. Bacteriol. 174, 6539-6547, 1992

A:Title: Sequence and characterization of the bacteriophage T4 comCalpha gene product, a protein involved in DNA replication

A:Reference number: A45731; MUID:93015705; PMID:1400206
A:Accession: D45731

A:Molecule type: DNA
A:Residues: 1-45 <SAN>

A:Cross-references: GB:M89919; NID:9215829; PDB:1AAA32486.1; PID:9215833
C:Genetics:
A:Gene: 39.2

A:Map position: 5.708-5.843

Query Match 43.1%; Score 25; DB 2; Length 45;
Best Local Similarity 44.4%; Pred. No. 5e+02;
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 YRLAIRRIA 9
|||:|:
Db 35 HRLAIRRIA 43

RESULT 15

T31016

hypothetical protein 52 - *Sulfolobus* sp. plasmid pNOB8

C:Species: *Sulfolobus* sp.

C:Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 02-Sep-2000

C:Accession: T31016

R:She, Q.; Phan, H.; Garrett, R.A.; Albers, S.V.; Steadman, K.M.; Zillig, W.

Extremophiles 2, 417-425, 1998

A:Title: Genetic profile of pNOB8 from *Sulfolobus*: the first conjugative plasmid from an archaeon

A:Reference number: Z20959; MUID:99044580; PMID:9827331
A:Accession: T31016

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA
A:Residues: 1-52 <SHE>

A:Cross-references: EMBL:AJ010405; NID:e1351926; PID:e1351938; PDB:1CAA09122.1
A:Experimental source: strain NOB8H2

C:Genetics:
A:Genome: plasmid pNOB8

Query Match 43.1%; Score 25; DB 2; Length 52;
Best Local Similarity 55.6%; Pred. No. 5.8e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 AIRRIALRY 12
|||:|:
Db 17 ALGNISLRY 25

Search completed: December 19, 2002, 16:22:56
Job time: 11.418 secs